IT PRO EN PYMES - NOTA #2 DE 5



$8,80 EN TODO EL PAIS

#29

Analizamos Vista Firewall

GRID COMPUTING

APLICACIONES PARA DISPOSITIVOS MÓVILES

VIRTUALIZACIÓN
XEN (OPEN SOURCE)
VIRIDIAN (MICROSOFT)
VMWARE

SEGURIDAD RFID RADIO FREQUENCY IDENTIFICATION

CLUSTERS MICROSOFT/LINUX

INNOVADORES IT



APRENDA CON LOS **MEJORES**

"Porqué la Virtualización es tan pupular"
Bill Hilf - General Manager of Platform Strategy de Microsoft

WWW.NEXWEB.COM.AR
ISSN 1668-5423
9 771668 542003 00029
Dist. Cap.: Vaccaro Sanchez y Cía. S.C. - Interior: DGP
Correo Argentino FRANQUEO A PAGAR Cta. 10185


☐ SOPORTA UNA PLATAFORMA DE GRAN ESCALABILIDAD
☐ GENERA CRECIMIENTO Y VALOR AGREGADO PARA LOS CLIENTES
☐ AUMENTA LA CAPACIDAD DE USUARIOS
☐ ACELERA Y SIMPLIFICA EL SERVICIO
☐ ES LINUX
☐ O WINDOWS SERVER
©2006. Microsoft Corporation. Todos los derechos reservados. Microsoft no ofrece ninguna garantía explícita o implícita. Este documento es sólo para fines informativos. Las marcas o nombres de productos aquí mencionados son marcas registradas o marcas de Microsoft Corporation en Estados Unidos u otros países. Otros productos o nombres de compañías mencionados en el presente, pueden ser marcas registradas de sus respectivos propietarios.

Microsoft



FIBERTEL AUMENTÓ 10 VECES LA CAPACIDAD DE USUARIOS DE SU SERVICIO FIBERWEB SIN INCREMENTAR SUS COSTOS, GRACIAS A WINDOWS SERVER 2003.

Fibertel lidera el mercado de acceso de alta velocidad y transmisión de datos en internet, con un gran potencial de crecimiento. La migración a Windows Server 2003 y el desarrollo de FiberWeb realizado por Ferengi en .NET permite ese incremento, posibilitando a los usuarios crear sus propias páginas web en forma sencilla y sin necesidad de conocimientos técnicos.

"Ser los líderes de acceso a Internet por banda ancha exige generar cada vez más y mejores servicios. Con Windows Server 2003, .NET y SQL Server podemos expandir nuestro servicio gratuito FiberWeb de 4.000 a 40.000 usuarios en tan sólo un año, optimizando nuestros resultados y reduciendo costos, ya que nuestros tiempos de administración se reducirán significativamente." - Fernando Casas, Gerente de Ingeniería de Cablevisión y Fibertel S.A.



Para mayor información de éste y otros casos, visite www.microsoft.com/argentina/hechos





DIRECTOR
- Dr. Carlos Osvaldo Rodriguez

PROPIETARIOS
- Editorial Poulbert S.R.L.

RESPONSABLE DE CONTENIDOS
- Dr. Carlos Osvaldo Rodríguez

DIRECTOR COMERCIAL
- Ulises Román Mauro
umauro@nexweb.com.ar

COORDINACIÓN EDITORIAL
- Alejandro Perakes
- Carlos Rodríguez

SENIOR SECURITY EDITOR
- Carlos Vaughn O'Connor

EDITORES TÉCNICOS
- María Delia Cardenal
- Thomas Hughes
redaccion@nexweb.com.ar

DISEÑO Y COMUNICACIÓN VISUAL
- DCV Esteban Báez
- Carlos Rodríguez Bontempi

DISTRIBUCIÓN
distribucion@nexweb.com.ar

ASISTENTE COMERCIAL
- Martín Guaglianone

SUSCRIPCIONES
- Maximiliano Sala
- Andrés Vázquez
suscripciones@nexweb.com.ar

PREIMPRESIÓN E IMPRESIÓN
IPESA Magallanes 1315. Cap. Fed.
Tel 4303-2305/10

DISTRIBUCIÓN
Distribución en Capital Federal y Gran Buenos Aires: Vaccaro, Sánchez y Cia. S. C. Moreno 794, Piso 9. C1091AAP- Capital Federal Argentina.
Distribuidora en Interior: DGP Distribuidora General de Publicaciones S.A. Alvarado 2118/56 1290 Capital Federal - Argentina
NEX IT Revista de Networking y Programación
Registro de la propiedad Intelectual en trámite leg número 3038 ISSN 1668-5423
Dirección: Av. Corrientes 531 P 1 C1043AAF - Capital Federal
Tel: +54 (11) 5031-2287

Si desea escribir para nosotros, enviar un e-mail a:
articulos@nexweb.com.ar



Nota del Editor

INTEROPERABILIDAD

Cuando discutimos qué destacar en la tapa de NEX #29, la palabra "interoperabilidad" surgió en varias ocasiones. Finalmente destacar el firewall de Windows Vista nos pareció también apropiado y nos decidimos por este último. Pero quien estudie los temas tratados en esta edición verá que varios artículos tratan de productos Microsoft que "inter-operan" con aquellos equivalentes Linux/Open Source.

Por ejemplo, para quienes necesiten realizar "number crunching" (cálculo de alto rendimiento o performance), Windows Compute Cluster Server 2003 (WCCS) es un nuevo producto de Microsoft y está diseñado para poder "interoperar" con Linux, el hasta ahora dueño y señor del High Performance Computing (HPC). Kyril Faenov, Director de Windows High Performance Computing, ha manifestado: "sabemos que HPC es heterogénea y con múltiples facetas... Aparte de permitir que partners puedan desarrollar productos para WCCS hemos puesto mucho énfasis en la integración con Linux. Por ejemplo tenemos una alianza con Plataforme Computing (vendor especializado en job scheduler) de modo que nuestros respectivos job schedulers puedan comunicarse el uno con el otro. WCCS 2003 también viene con Microsoft Message Passing Interface, basada en el Standard MPI2 para poder portar aplicaciones paralelas de modo sencillo".

En el mundo de servidores Windows el R2 viene con una serie de tecnologías que le permiten interactuar con máquinas Unix. Solo algunos ejemplos: Directory Services viene nativa con autenticación Kerberos, SUA (Subsystem for Unix-based Applications), que permite una administración a través de scripts ya sea sobre Unix o Windows. Bajo SUA las aplicaciones bajo Unix pueden recompilarse y correrse en Windows. NFS (Network File System) y sincronización de passwords son otros dos ejemplos.

Pero quizá donde esta interoperabilidad se hace más destacable es en lo relacionado a las nuevas tecnologías de virtualización basadas en el concepto de hypervisor (Viridian y Carmine). Pero aún con productos actuales, hoy ya es posible correr máquinas virtuales Linux sobre "Virtual Server". No deje de leer los artículos sobre el tema si desea entender la alianza de Microsoft y XenSource y conocer qué se hace en la UBA sobre clusters y Grid Computing.

En próximas ediciones de NEX trataremos otras herramientas de interoperabilidad propuestas por Microsoft. Fundamentalmente las que apuntan a la administración de sistemas heterogéneos: System Center Operations Manager (OPS Manager, antes llamado MOM) y System Center Configuration Manages (SCCM, antes SMS).

Como en todos los NEXs el espectro de artículos es amplio. NEX #29 se compone además de artículos de networking, desarrollo para dispositivos móviles, seguridad en la tecnología RFID y cómo manejar el correo no deseado con Exchange entre otros.

Sé que disfrutarán los artículos y, como siempre, no dejen de contactarnos a redaccion@nexweb.com.ar

SUMARIO

Pag.38

IT Pro en PyMEs Nota 2

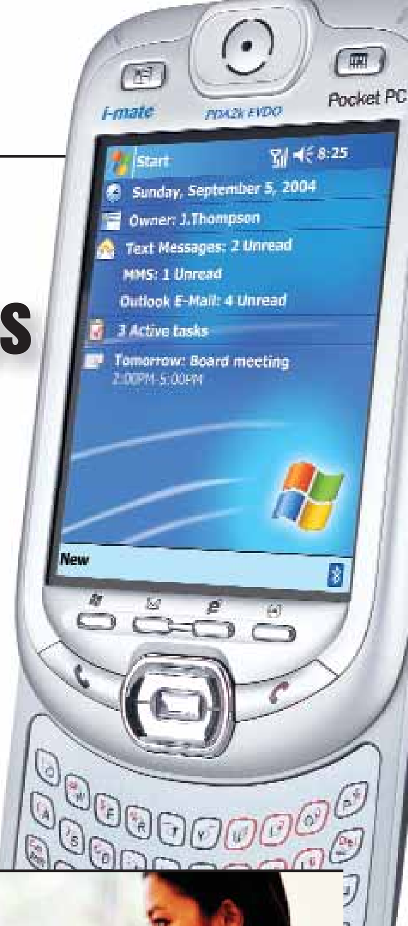
SAM es, además de una herramienta, una guía de buenas prácticas, la cual permite al departamento de IT lograr todos los objetivos en cuanto a la administración de activos de software se refiere.

Arquitecturas para el desarrollo de

Aplicaciones Móviles

La telefonía celular brinda una amplia variedad de arquitecturas de comunicación para el desarrollo de aplicaciones móviles. En esta nota analizamos algunas de ellas y los diferentes tipos de aplicaciones posibles.

Pag.12



Fundamentos de Networking

La Capa Física

La primera capa del modelo OSI se encarga de recibir los datos de las capas superiores y convertirlos en un flujo de bits, para luego enviarlos hacia el destino. Las funciones de esta capa están íntimamente relacionadas con los parámetros del medio físico.

Pag.32

05 **Nota del Editor**

07 **Eventos**

08 **Windows Vista Firewall**
Conozca uno de los componentes más importantes que mejora la seguridad de nuestro equipo.

12 **Aplicaciones Móviles**
Analizamos algunas de las arquitecturas de comunicación para las aplicaciones móviles.

20 **Correo NO Deseado**
Exploramos la experiencia de una pequeña empresa que logró evitar que el spam invadiera su bandeja de entrada.

24 **Radio Frecuency Identification**
La revolución RFID y sus riesgos asociados.

30 **Ethernet in the First Mile**

32 **La Capa Física**
La Capa Física opera solamente a nivel de 1s y 0s y es responsable por la transmisión de la información, como un flujo bits.

38 **IT PRO en PyMEs**
Cómo auditar nuestros activos de software a través del SAM: Software Asset Management.

42 **Terminal Server**
Cuáles son los nuevos componentes de la versión Beta 2 que permiten hacer de Terminal Server una solución de acceso a aplicaciones propia.

46 **ITIL: ¿En qué fase estamos?**

48 **Cluster Beowulf bajo Linux**
Entérese de qué es un cluster Beowulf, cómo funciona, cómo se instala y qué beneficios nos brinda.

54 **Windows Compute Cluster Server 2003**
Cuáles son los requerimientos, tecnologías y beneficios del Windows Compute Cluster Server 2003.

58 **Clusters y Grid**
Conozca cómo se aplica Hight Performance Computing en la investigación universitaria y qué proyectos posibilita.

62 **openMosix y Condor**
Cuáles son los tipos de clusters existentes y cómo operan openMosix y Condor para darnos una solución a los problemas de balance de carga.

66 **Virtualización: Viridian + Carmine**
A través de la percepción de Bill Hilf, General Manager of Platform Strategy de Microsoft, conozca porqué la virtualización es tan popular.

68 **Business Intelligence y Data Mining**
Cuáles son los componentes de Business Intelligence y Data Mining y cómo nos ayudan a anticipar nuevas oportunidades.

74 **.NET 2.0 Framework**
Cuáles son las nuevas funcionalidades que mejoran el rendimiento del Framework 2.0 con respecto a los anteriores.

78 **Futuro Laboral IT**

80 **Novedades de Productos y Empresas**

82 **Breves**

Aclaración: *Por error de edición, en el número anterior de NEX IT Specialist, se publicó equivocadamente el autor de una nota. Link Aggregation fue escrita por Miguel F. Lattanzi, Ing. en Telecomunicaciones.*

6 años de Snoop Consulting

La empresa especializada en servicios sobre nuevas tecnologías de información cumplió seis años en el mercado y lo festejó a fines de agosto con un evento en Opera Bay. El tema de la fiesta fue "Una noche en la Opera" y por tal motivo contaron con la participación del cuarteto de cuerdas "Del Bosque" y del grupo Opera Prima Rock. Además la animación estuvo a cargo del cantante Manuel Wirtz quién divirtió a más de 500 personas entre empresarios del sector tecnológico, ejecutivos de informática, empleados y amigos de Snoop Consulting.

"Esto que tenemos hoy es el resultado de seis años de un trabajo serio y constante, que nos ha permitido casi duplicar nuestro tamaño en cada año de vida", señaló Gustavo Guaragna, Presidente y CEO de Snoop. "Nuestra filosofía ha sido construida sobre bases firmes: la formación de un sólido grupo profesional que crece y se capacita constantemente, una importante inversión en Investigación y Desarrollo y una gran agilidad para adecuarnos a las necesidades del mercado y a nuestro propio proceso de crecimiento."

Por su parte, Sergio Candelo, Director de Operaciones de la compañía, elogió la celeridad de trabajo de la AFIP, gracias a la cual Snoop, beneficiada por la "Ley del Software" a partir de julio pasado, ya ha comenzado a percibir los beneficios tributarios otorgados por la misma.

Acerca de Snoop Consulting

Snoop Consulting es la empresa especializada en brindar soluciones sobre nuevas tecnologías de la información que permitan a sus clientes colocarse a la vanguardia de los negocios minimizando los riesgos inherentes al cambio mediante una adecuada transferencia tecnológica. Snoop provee servicios de valor agregado en tecnologías de desarrollo orientadas a objetos (Java y .NET), integración de aplicaciones y SOA (Arquitectura Orientada a Servicios), High Performance Computing y Tuning de Aplicaciones, Data Mining, Visualización de Datos y Usabilidad de interfaces de usuario. Basa sus soluciones en plataformas Oracle, Microsoft, IBM, BEA, Red Hat Linux y Open Source. ●

Más información en: *http://www.snoopconsulting.com*







CALENDARIO DE EVENTOS IT EN ARGENTINA PARA EL 2006

Fecha		Informes
	OCTUBRE	
3 al 6	**EXPO COMM** - La Rural, Predio Ferial de Buenos Aires.	www.expocomm.com.ar
9 a 13	**InfoSecurity Week Buenos Aires** - Hotel Hilton Buenos Aires	www.infosecurityonline.org
13 a 15	**6ta Jornadas Regionales de Software Libre** - Universidad de Mendoza	http://jornadas.lugmen.org.ar
18	**Snoop Update ´06** - Paseo La Plaza	www.snoopconsulting.com/update06/
18	**Business Continuity Planning** - Buenos Aires Sheraton Hotel	www.cybsec.com/capacitacion
26 y 27	**Jornadas Nacionales AGSI 2006** - Hotel Sheraton Libertador	www.worktec.com.ar
	NOVIEMBRE	
1	**Jornadas Trabajo IT 2** - Sheraton Libertador.	www.trabajoit.com.ar
1	**2do Congreso Nacional de Estudiantes de Sistema y Tecnologías de la Información** - A confirmar	www.worktec.com.ar - info@worktec.com.ar
2 al 5	**AES - Argentina Electronic Show** - La Rural, Predio Ferial de Buenos Aires.	www.aeshow.com.ar/es_services_contact_us
10 y 11	**CaFeCONF 2006** - UADE, Buenos Aires	www.cafeconf.org
13 al 16	**Cisco Networkers Solution Forum 2006** - Hotel Hilton Buenos Aires	www.cisco.com/ar/networkers
14 y 15	**Consecri-Consetic 2006** - Sheraton Libertador.	www.consetic.com.ar / www.consecri.com.ar

Si desea ver su evento IT publicado en esta sección, por favor háganos llegar la información respectiva a: eventos@nexweb.com.ar

Security Summit

El 12 de septiembre se realizó el "Security Summit", evento organizado por Axis Group en el Hotel Hilton., donde se dieron a conocer las mejores prácticas del sector sobre soluciones tecnológicas para contribuir al negocio y medir el retorno de la inversión.

Los encargados de las disertaciones fueron Claudio Avin, Señor Security Consultant de Symantec; Fabián Domínguez, Gerente de Desarrollo de negocios para Tecnologías Avanzadas de Cisco System quien habló sobre cómo convertir una red actual en la mejor defensa contra amenazas; y Daniel Levi, regional IT Manager de Microsoft quien dio a conocer una visión general de sus prácticas internas.

Windows Vista Firewall

Como frecuentemente comento, la tendencia en lo que respecta a seguridad de los productos Microsoft ha cambiado desde hace ya un tiempo. Es por eso que con sólo instalarlos ya obtenemos una configuración segura por defecto. Y el Windows Firewall no es la excepción.

A partir del SP1 de 2003 y el SP2 de XP ya podemos relajarnos un poco a la hora de conectarnos a Internet, o incluso a una gran red corporativa. Ya que por default tenemos un Firewall (accesible desde el panel de control) con una regla que deniega todo tráfico entrante, en donde sólo se permite el ingreso de una respuesta a una conexión generada desde nuestro host o una conexión entrante que esté expresamente exceptuada de la regla general que deniega todo tráfico entrante. Por ejemplo, si publicara un FTP en mi Windows XP SP2 con el servicio de Firewall levantado. Para permitir el acceso debería crear una excepción a la regla de tráfico entrante que existe por default especificando ese protocolo. No así para las conexiones salientes ya que no se restringe este tipo de tráfico pudiendo originar cualquier tipo de conexión hacia cualquier lado y permitiendo la respuesta. Entonces uno pensaría "la idea es defender nuestro equipo del mundo, y no al mundo de nuestro equipo" lo cual tiene su lógica, pero a un administrador de red o usuario avanzado a veces se le hace vital poder tener un control sobre las conexiones que puede y no puede realizar un cliente.

Algunos alumnos en mis clases me han consultado:
"Si se metió un virus en mi máquina que utiliza conexiones hacia todas las máquinas de mi Lan en un determinado rango de puertos y protocolo, ¿cómo puedo hacer para bloquear esas conexiones?"
"Administro una empresa realmente grande y hay empleados que jugando a ser hackers utilizan programas de administración remota y troyanos para conectarse con otros hosts de mi red interna hacia puertos conocidos desde un determinado rango de puertos, ¿cómo logro impedir esas conexiones?"
Y mi respuesta, luego de darles un sermón sobre antivirus, auditoría, etc., era:
"Para controlar conexiones entrantes y salientes en forma centralizada (sin la necesidad de productos de terceros) podés configurar IPSec por políticas de dominio".
Lo cual no siempre los tranquiliza, por el desconocimiento de las políticas y/o dificultad de su uso. Por lo que nunca falta uno que vuelva a la carga y pregunte si no hay un método más sencillo que directamente bloquee la conexión que se intenta hacer (por ejemplo que ninguna máquina pueda conectar al puerto X desde el puerto Y; o mejor aún que no permita la conexión de determinada aplicación o servicio -como un virus-). Y en ese punto, con una gota gorda naciendo en mi frente estaba obligado a responder que no se pueden controlar las conexiones salientes a través del Firewall integrado, y menos con tanto nivel de detalle.
Afortunadamente, ya no me encuentro en esa difícil situación en mis cursos, porque el Firewall de Windows Vista ya nos da respuestas complacientes a todo esto.

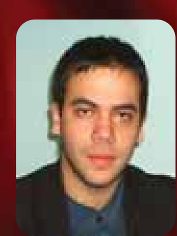
Lo Nuevo

A simple vista podemos ver en el panel de control la misma consola de Firewall que el anterior con los mismo seteos por default (figura 1). Pero ahora se agrego una nueva MMC llamada "Windows Firewall with Advanced Security" (figura 2).



Otra vuelta de tuerca le ha dado Microsoft a un producto que nació con el SP2 de Windows XP y el SP1 de Windows 2003. Nos referimos al Firewall que trae integrado Windows Vista, uno de los componentes vitales que mejora sustancialmente la seguridad de nuestro equipo.

Por: **Ariel Giarratana**
Senior Trainer
MCSE + Security



Esta consola nos permite realizar seteos avanzados tanto para conexiones entrantes, como salientes (cosa que no se podía en las versiones anteriores). Siguiendo la misma política para conexiones entrantes "no permito ingresar nada, salvo las excepciones". Y para las salientes es al revés "Dejo salir todo, salvo las excepciones". De todas maneras todo esto es configurable (figura 3).
Desde aquí podremos configurar los perfiles, de acuerdo a nuestro tipo de conexión. Desde esta MMC podremos configurar tanto nuestro Firewall como el de una máquina remota. También es posible realizar los seteos en forma centralizada a través de Políticas de Grupo.

Para configurarlo a través de group policies hay que ir a:
"Computer Configuration\Windows Settings\Security Settings\Windows Firewall with Advanced Security" en la consola de política de grupo. Estos seteos van a ser tomados por el nuevo Firewall, por lo que las computadoras con XP SP2 ó 2003 SP1 van a ignorar la mayoría.
Los seteos para el firewall "anterior" se configuran de "Computer Configuration\Administrative Templates\Network\ Network Connections\Windows Firewall" (Vista tomaría también estos seteos).
También desde línea de comandos podemos realizar seteos avanzados con el comando:

netsh advfirewall

opción nueva del *netsh* que no existía en versiones anteriores.

Firewall & IPsec sean unidos

Antiguamente las reglas de IPsec y de firewall estaban separadas. Lo cual no sólo hacía más complicada la administración, sino que daba lugar a que se solapen reglas que afecten al mismo tipo de conexión. A partir de ahora todo esto está integrado en la misma consola y el uso de las políticas de IPsec se simplifica significativamente.

< Fig.1 - Panel de Control

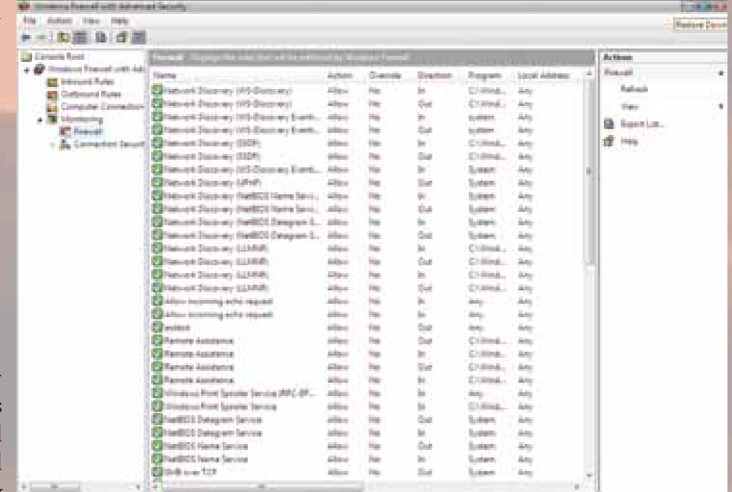
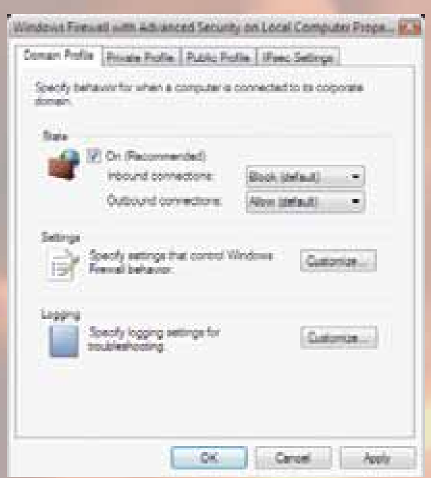


Fig.2 > Windows Firewall with Advanced Security



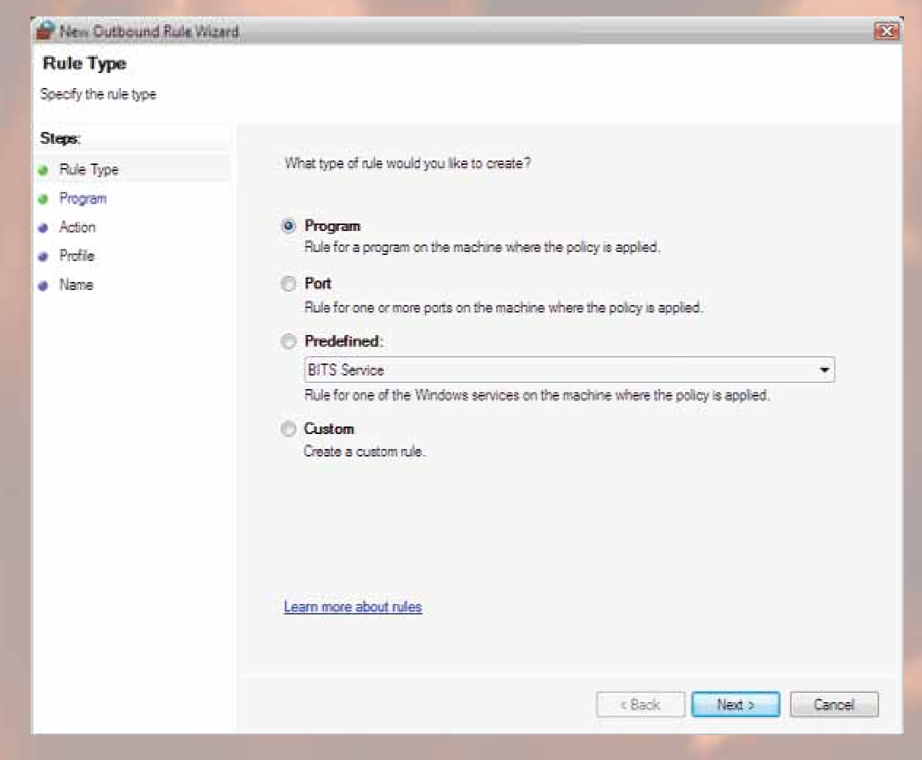
^ Fig.3 - Configuración de Perfiles

Windows Vista Firewall

Esta consola de configuración avanzada nos da la posibilidad de definir varias condiciones para las reglas:

Por cuentas y Grupos de Active Directory
Para las reglas de IPsec se podrán especificar las cuentas de usuario, de computadora y grupos que estén autorizados para iniciar la comunicación. Por ejemplo se puede definir qué conexiones a ciertos servidores deben estar protegidas y sólo pueden ser originadas desde determinadas computadoras o cuentas de usuario.

Por dirección IP de origen y destino
Supongamos que quiero que ciertas IP's se conecten a un grupo de servidores. Puedo generar una excepción de tráfico saliente en donde las IP's que desee no puedan conectarse a las IP's de los servidores. También se pueden configurar Ip's de origen para las conexiones entrantes a nuestro host. También existen direcciones predefinidas para indicar en la regla.



• *Default gateway, WINS servers, DHCP servers, DNS servers*
Estas direcciones son mapeadas dinámicamente del default gateway, WINS, DHCP, y DNS servers que tenga configurado el host.

• *Local subnet*
Esta subred halla el ID de red dinámicamente de la dirección IP y submáscara de red que tengamos configuradas.

Por número de protocolo IP
En el viejo Firewall se podía definir reglas basadas sólo para tráfico TCP o UDP. En el nuevo se puede elegir el protocolo predefinido por nombre, o manualmente tipear el valor del protocolo IP.

Por puerto de origen y destino
Con el viejo firewall, se podía especificar sólo el puerto TCP o UDP de destino para la conexión entrante. Ahora se podrá definir el puerto de origen y destino tanto para las conexiones entrantes como para las salientes

Por ejemplo si un virus, troyano, o lo que sea intenta realizar conexiones a puertos conocidos desde un determinado puerto o rango de puertos, puedo especificar que dicha conexión saliente esté exceptuada. O para las entrantes, por ejemplo, permitir conexiones sólo desde los puertos que correspondan.

Para todos o varios puertos
Antes sólo se podía especificar de a un puerto por excepción de regla. Ahora tendremos la opción de "Todo tráfico TCP" o "Todo tráfico UDP", o la opción de especificar varios puertos delimitados por comas. Por ejemplo cuando los puertos de origen son del 3000 al 3003 se debería poner 3000,3001,3002,3003.

Por tipo de Interface
Antes las reglas aplicaban a todas las interfaces. Ahora se puede especificar a cuáles. Se incluyen en forma predefinida LAN, remote access o Wireless.

Por servicios
Se podrá especificar si la regla aplica a todos los procesos, solo a servicios, o al servicio especificado (indicando su nombre y path).

Por Tipo y Código de tráfico ICMP
Antes estaban predefinidos los tipos de excepciones que podíamos crear para las conexiones ICMP. Ahora podremos especificar el tipo y código de ICMP.

Consola avanzada
Volviendo a la consola avanzada (figura 2) analicemos cómo está organizada.
Para el que haya utilizado ISA Server 2004 o 2006 tal vez le suene un tanto familiar la organización de la consola.
Hay 3 paneles:
• En el de la izquierda tenemos 4 ramas:
- *Inbound Rules*
- *Outbound Rules*
- *Connections Security Rules*
- *Monitoring*
• En el del medio la configuración actual
• En el de la derecha las acciones que se pueden realizar de acuerdo a la rama en la que nos encontremos situados del panel izquierdo.
Si en el panel izquierdo hacemos botón derecho sobre Windows Firewall, propiedades, podremos configurar las reglas por default para los diferentes perfiles (si estamos conectados a un dominio, a una red privada, o a una red pública) y el logueo.

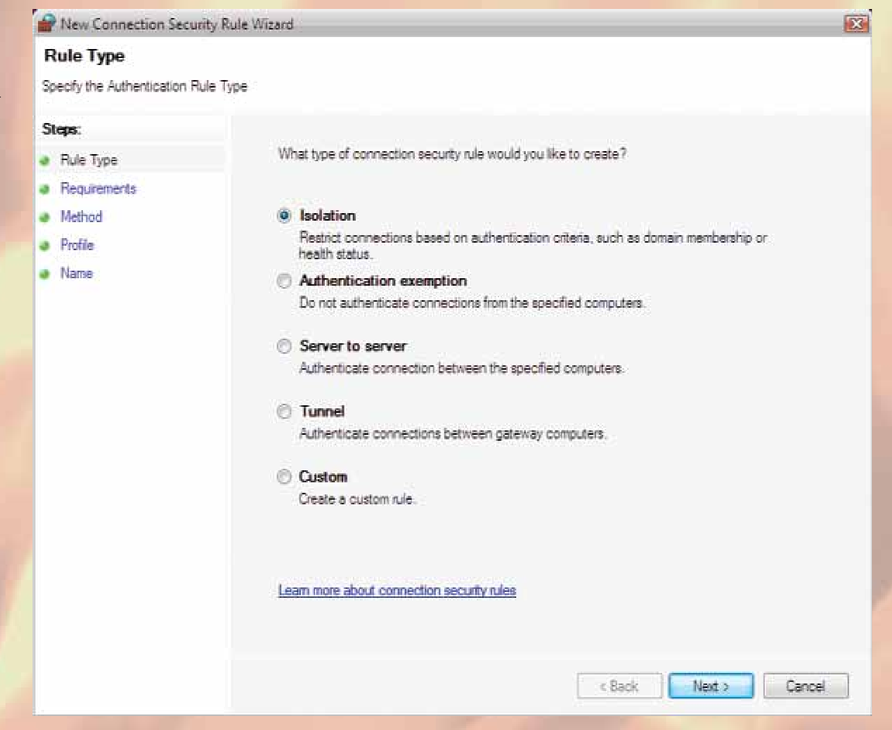
Excepción para tráfico entrante

Para crear una regla para tráfico entrante nos situamos sobre Inbound rules y hacemos click en new rule sobre el panel de Acciones (figura 4).



< Fig.4 - Rule Type

Connection Security Rule - Fig.5 >



Gracias a que las reglas de IPsec y de firewall están integradas en una misma consola nos da la posibilidad de definir varias condiciones para las reglas Por cuentas y Grupos de Active Directory, Por dirección IP de origen y destino, Por número de protocolo IP, Por puerto de origen y destino, Para todos o varios puertos, Por tipo de Interface, Por servicios y por Por Tipo y Código de tráfico ICMP.

Básicamente hay que realizar una serie de pasos. Podemos especificar una regla para tráfico entrante basada en un programa (especificando el ejecutable), para puerto (basado en el puerto TCP o UDP de la conexión), reglas predefinidas, o una customizada cuando ninguna de las anteriores cumpla con nuestros requerimientos. También debemos configurar una acción (Permitir, Denegar o Proteger), un perfil (Domain, Private y Public) y un nombre.
Luego de haber creado la regla, se podrá editarla seleccionando la regla y marcando propiedades en el panel de acciones.
Nos encontraremos con varias solapas para poder configurar la regla a nuestro gusto.

- **General:** en donde podremos configurar el nombre y la acción.
- **Programs and Services:** los programas y servicios a los cuales le aplica la regla.
- **User and Computers (inbound) or Computers (outbound):** si la acción de la regla es "permitir sólo conexiones seguras", las computadoras y usuarios definidos podrán realizar la conexión.
- **Protocols and Ports:** se define el protocolo IP, origen y destino para las conexiones TCP y UDP , y los seteos ICMP.
- **Scope:** direcciones de origen y de destino.
- **Advanced:** perfiles o tipos de interfaces al los cuales la regla aplica.

Configurar una conexión saliente es prácticamente igual.

Connection Security Rules

Ahora las reglas de IPsec se pueden configurar desde "Connection Security Rules" (figura 5). El cual nos ofrece un Wizrad más sencillo e intuitivo para generar conexiones seguras (encriptación e integridad).

Otra vez, debemos pasar por un Wizard en donde definiremos:

- **Tipo de conexión:** Isolation, Authentication Exception, Server to Server, Túnel o Custom
- **Requerimientos:** si la seguridad será obligatoria o no -Require o request-
- **Método de Autenticación:** Kerberos, Certificados, Preshared Key
- **Perfil y nombre**

Y nuevamente podremos modificar las reglas yendo a las propiedades de las mismas, en donde nos encontraremos con las siguientes solapas:

- **General:** para configurar el Nombre de la regla y si la regla está habilitada.
- **Computers:** el set de computadoras, por dirección IP, para las cuales se realiza el tráfico protegido.
- **Authentication:** en donde se juntan requerimiento de seguridad y autenticación.
- **Avanzado:** perfiles, tipos de interfaces a las cuales aplica y los endpoints si es se trata de un tunel IPSec .

Conclusión

Hay que ser conciente de que con todas estas novedades estamos agregando un nuevo nivel de seguridad a nuestra red.

La inseguridad no está de nuestro Proxy para afuera, también está adentro. Por eso es que me parece excelente que tengamos una manera tan sencilla de configurar en forma centralizada, qué pueden y qué no pueden hacer nuestros clientes en la red y al mismo tiempo asegurar qué pueden aceptar y qué no. Y además darnos el lujo de poder ponernos en exquisitos en cuanto a las reglas que queremos para nuestros clientes. Y todo esto en forma integrada y sin la necesidad de pagar por software extra. ●

Arquitecturas para el desarrollo de Aplicaciones Móviles

Por **Gabriel Holzmann**
Presidente de Behind Systems

Introducción

Si tenemos presente que el celular es un hardware capaz de transmitir voz y datos en diferentes formatos: mensajes de texto, imágenes, video; y que adicionalmente puede brindar información de posicionamiento mediante GPS o triangulación, entendemos que estamos ante uno de los más completos circuitos de comunicación e información del mercado.
Si adicionalmente, concluimos que la mayoría de las personas, para uso individual o empresarial, por comunicación, hobby o trabajo, lleva a todos lados al menos un celular, entendemos que el mercado de los tipos de aplicaciones posibles, y la cantidad de usuarios potenciales son enormes.
Sin embargo, encontramos un déficit en la oferta de soluciones tendientes a resolver negocios empresariales, dejando casi la totalidad del mercado al consumo masivo de la recreación personal, por ejemplo los mensajes de texto y las trivias.
Es la intención de esta nota, presentarles la arquitectura de las aplicaciones SMS (mensajes de texto y trivias) como punto principal, y la arquitectura de las aplicaciones remotas utilizando WAP/GPRS (navegación web desde el celular) de forma complementaria.
Adicionalmente, nombraremos otras posibilidades técnicas para el desarrollo de aplicaciones embebidas en el celular.

Soluciones SMS

En esta sección, describiremos los protocolos necesarios, centrándose en la capa de transferencia de mensajes, la interfaz entre las aplicaciones y el servicio SMS, utilizando un módem GSM para tener acceso a ellas, y mostraremos cómo es posible controlar este módem mediante los comandos AT y AT+.

Introducción a GSM

GSM (Global System for Mobile Communication) es un sistema de comunicaciones móviles que ofrece los siguientes servicios:
- Transmisión/recepción de voz
- Transmisión/recepción de datos
- Envío/recepción de mensajes cortos (Sort Message Service, SMS)

Servicio SMS

El servicio SMS permite transferir un mensaje de texto entre una estación móvil (MS) y otra entidad (SME) a través de un centro de servicio (SC) (Figura 1).
El servicio final ofrecido es una comunicación extremo-extremo entre la estación móvil (MS) y la entidad (SME). La entidad puede ser otra estación móvil o puede estar situado en una red fija. En el caso de envío de un mensaje entre dos móviles, ambas partes son estaciones móviles. Cuando se envía un mensaje para solicitar algún tipo de servicio (o realizar alguna votación, sobre todo en los concursos de la TV, que ahora están tan de moda), un extremo es una estación móvil y la otra es un servidor que atiende las peticiones (o anota los votos).
El servicio SMS se divide en dos servicios Básicos (Figura 2):
1. SM MT (Short Message Mobile Terminated Point-to-Point). Servicio de entrega de un mensaje desde el SC hasta una MS, obteniéndose un informe sobre lo ocurrido.
2. SM MO (Short Message Mobile Originated

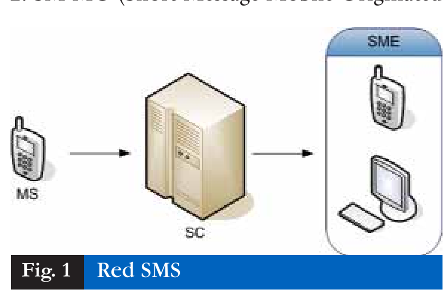


Fig. 1 Red SMS

La telefonía celular brinda una amplia variedad de arquitecturas de comunicación para el desarrollo de aplicaciones móviles. En esta nota analizamos algunas de ellas y los diferentes tipos de aplicaciones posibles.

Point-to-Point). Servicio de envío de un mensaje desde una MS hasta un SC, obteniéndose un informe sobre lo ocurrido.

Arquitectura básica de la red SMS

La estructura básica de la red para el servicio SMS se compone de las siguientes entidades:
- MS: Estación móvil
- MSC: Centro de conmutación
- SMS-GMSC: MSC pasarela para el servicio de mensajes cortos (Servicio SM MT)
- SMS-IWMSC: MSC de interconexión entre PLMN (Public Land Mobile Network) y el SC

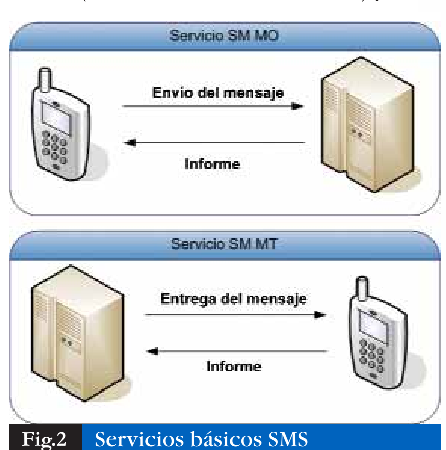


Fig.2 Servicios básicos SMS

(Servicio SM MO)
- SC: Centro de Servicio
- HLR (Home Location Register), VLR (Visitor Location Register)

Para la descripción detallada de la arquitectura, se utiliza un modelo de capas, en el que cada capa o nivel proporciona un servicio a la capa superior, y este servicio se implementa mediante el protocolo correspondiente.

La arquitectura se divide en 4 capas (Figura 3):

- SM-AL (Short Message Aplication Layer): Nivel de aplicación.
- SM-TL (Short Message Transfer Layer): Nivel de transferencia. Servicio de transferencia de un mensaje corto entre una MS y un SC (en ambos sentidos) y obtención de los correspondientes informes sobre el resultado de la transmisión. Este servicio hace abstracción de los detalles internos de la red, permitiendo que el nivel de aplicación pueda intercambiar mensajes.
- SM-RL (Short Message Relay Layer): Nivel de repetición. Proporciona un servicio al nivel de transferencia que le permite enviar TPDU (Transfer Protocol Data Units) a su entidad gemela.
- SM-LL (Short Message Lower Layers): Niveles inferiores.

Utilización de los servicios SMS

Actualmente están apareciendo gran cantidad de servicios basados en mensajes cortos. Además de ser usados para enviar mensajes de texto entre personas, se están ofreciendo otros servicios como ser:

- **Correo electrónico.** Se convierte un mensaje de correo en SMS y un mensaje SMS en mensaje de correo, con las consiguientes generaciones/eliminaciones de cabeceras de mensaje.
- **Distribución de mensajes SMS.** De la misma forma en la que funcionan las listas de correo electrónico, en las que un mensaje es reenviado a los suscriptores de dicha lista, en la lista de distribución de mensajes SMS se permite el mantenimiento (alta/baja/consulta) de suscriptores y, de esta forma, se le envía al resto de la lista los mensajes que no son comandos de actuación sobre la propia pasarela.
- **Recepción de alarmas de los sistemas de monitorización de servicios.** Aplicaciones generan avisos cuando se alcanzan ciertos eventos. Estos avisos pueden ser encaminados mediante mensaje SMS dependiendo de su importancia, para que ciertas personas sean avisadas inmediatamente.
- **Transporte de contenidos web.** El SMS es utilizado como paquete de transporte para hacer llegar desde el móvil al servidor la petición de una página web y desde el servidor al móvil el contenido de dicha página una vez "filtrada" para eliminar imágenes, tags html, cabeceras de página, etc.
- **Concursos de preguntas y respuestas.** Ante una solicitud desde el teléfono móvil para comenzar el concurso, la pasarela envía mensajes con preguntas, recibe respuestas y mantiene un contador de resultados para cada participante, de manera de generar las clasificaciones.
- **Sistemas de seguimiento de flotas de vehículos.** Un teléfono celular unido a un módulo GPS permite enviar información acerca de la posición exacta del portador del teléfono, de manera que para flotas de vehículos se reciben sus notificaciones de posición y ésta actualiza una base de datos consultable por otras aplicaciones que pueden mostrar la situación de cada elemento de la flota.
- **Notificación de estado de dispositivos aislados.** Máquinas de vending, dispositivos de control de temperatura y presencia, etc. que se encuentren aislados y sin comunicación con una red mediante enlace físico pueden hacer uso de los mensajes SMS para recibir ordenes y para notificar su estado (queda poca bebida, la temperatura ha superado los 45 grados, etc.). Normalmente esta comunicación se realiza sin intervención manual, por lo que realmente se conectan equipos automáticos en ambos lados. Es lo que se conoce como M2M (machine to machine).
- **Banca GSM, Notificaciones Académicas, etc.**

SME
SC
SMS-GMSC
SMS-IWMSC
MSC
MS
SM-AL
SM-TL
SM-RL
SM-LL

Fig.3 Arquitectura de red SMS

Para poder ofrecer estos servicios es necesario diseñar software y hardware que pueda acceder a los servicios SMS. Esto se puede conseguir de varias maneras:

- Algunos teléfonos se pueden conectar directamente a un PC y mediante un software propietario se puede acceder a los datos de móvil (agenda, tarjeta SIM, etc), así como enviar y recibir mensajes SMS. El principal problema de esta solución es que no es abierta, y los fabricantes no proporcionan suficiente información como para poder realizar aplicaciones con ellos. Es necesario realizar ingeniería inversa.



- Utilización de un MÓDEM GSM. Esta es la opción por la cual nos centraremos porque nos permite escribir nuestras propias aplicaciones con frameworks tales como Microsoft .NET

Mediante un MÓDEM GSM podemos conectar cualquier sistema digital a la red GSM, no sólo para enviar mensajes SMS sino también para transmitir datos.
Existen dos tipos de MODEMS, según la aplicación que queramos realizar:
- MODEMS para circuito impreso: Son Modems de reducido tamaño y perfectamente apantallados que están preparados para ser incorporados dentro de un circuito impreso y que permiten desarrollar un hardware específico y que no depende de un PC.
- MODEMS para PC: Tienen un tamaño también bastante reducido y disponen de un cable estándar para conectarse a la PC. Son muy útiles para permitir que desde cualquier ordenador de una intranet se puedan enviar mensajes SMS.

Interfase con módems GSM

Es en esta parte de la arquitectura donde entra en juego nuestra aplicación personalizada para responder a los requerimientos del negocio. Supongamos que necesitamos implementar un servicio de alerta a los operadores de una empresa. La red GSM y los celulares de los operarios ya los tenemos. Ahora nos queda incorporar un módem GSM a un servidor de aplicaciones y desarrollar la solución de software encargada de interactuar con el módem y con los sistemas internos de la empresa. ¿Cómo podemos lograr que nuestro framework SMS se comunique con el modem GSM?
Los modems GSM no sólo se comportan de forma muy parecida a un modem normal, permitiendo el intercambio de datos con otro modem y utilizándose los comandos AT originales, sino que incluyen muchas más características. Son como pequeños teléfonos móviles, que incluyen su propia tarjeta SIM para poder funcionar y por tanto permiten gestionar la base de datos de teléfonos, la lista de los mensajes SMS recibidos, enviar mensajes SMS y configurar diversos parámetros.
Para tener acceso a todos esos servicios, y dado que los comandos AT estaban muy extendidos y muy estandarizados, se ha realizado una ampliación, añadiéndose nuevos comandos. Estos nuevos comandos comienzan por las letras AT+, y se denominan comandos AT+. Para una lista de los comandos AT+ vea el recuadro “Comandos AT+”.
Para probar los comandos puede utilizar el programa Hyperterminal de Windows o desarrollar una pequeña aplicación.

Rendimiento

La utilización de un módem GSM y la realización de una aplicación personalizada, contienen la ventaja de mantener toda la lógica dentro de la empresa. Sin embargo, cuando pensamos en términos de rendimiento, nos encontramos con la limitación de la cantidad de mensajes que puede procesar un módem GSM por minuto.
Si tomamos en promedio el proceso de un mensaje cada cuatro segundos, nos encontramos con un techo de 900 mensajes por hora por MÓDEM. Es cierto que podemos conectar más de un MÓDEM al servidor, pero igualmente nos encontramos con soluciones no muy escalables y difíciles de mantener.
Para la gran cantidad de nuestras aplicaciones, este valor es más que suficiente, sin embargo, para aquellas de uso masivo donde en una hora pico podrían recibirse 100.000 o más mensajes, deberíamos pensar en otra arquitectura.
Es aquí donde surgen dos posibilidades adicionales:
- La utilización de un gateway como intermediario.
- Un enlace directo con las TELCOs.

Utilización de un gateway como intermediario

En este caso nos referimos a contratar los servicios de una empresa dedicada a brindar servicios de mensajería y que recibe comúnmente el nombre de gateway SMS aunque también brinda otros servicios.
En el país hay 4 ó 5 gateways importantes que manejan una gran porción del mercado de aplicaciones SMS.
Nosotros podríamos instalar un gateway SMS como Kannel, Alamin u otros más comerciales, pero deberíamos dedicarnos a un mantenimiento continuo. De todas formas, un análisis más detallado de los mismos quedará para futuras entregas.
El punto es que, mediante la contratación de un gateway SMS, resolvemos los problemas aparentes de escalabilidad para acceder a un mayor promedio de mensajes por hora.
Los gateways tienen pre-contratos realizados con las TELCOs que nos permiten exponer nuestras soluciones de negocio para ser utilizadas por cualquier usuario, independientemente de la empresa de telefonía celular.
La aplicación que deberíamos desarrollar en por ejemplo Microsoft .NET, podría hostearse dentro de las instalaciones del gateway o en las instalaciones de la empresa. En ambos casos, la comunicación entre el gateway y nuestra aplicación hosteada y/o nuestra aplicación hosteada y el proxy instalado dentro de la empresa, se realizaría mediante protocolos de red como por ejemplo SMPP para el envío y recepción de mensajes SMS.

Enlace directo con las TELCOs

Cuando la utilización de la empresa gateway como intermediario, por cuestiones de rendimiento económico no conviene al negocio, siempre existe la posibilidad de habilitar un enlace directo entre las TELCOs y la empresa. En este caso, se puede generar un enlace seguro y utilizar SMPP para intercambiar los mensajes entre ambas entidades.
Esta solución es tal vez la de mayor escalabilidad. Sin embargo esa escalabilidad debe ser mantenida por la empresa, por cuanto tal vez

Comandos AT+

Algunos de los comandos AT+ implementados en los modems GSM para tener una idea de lo que se puede controlar a través del modem, aunque existen una mayor cantidad, son los siguientes:

1) Comandos generales
I) ***AT+CGMI:*** Identificación del fabricante
II) ***AT+CGSN:*** Obtener número de serie
III) ***AT+CIMI:*** Obtener el IMSI.
IV) ***AT+CPAS:*** Leer estado del modem

2) Comandos del servicio de red
I) ***AT+CSQ:*** Obtener calidad de la señal
II) ***AT+COPS:*** Selección de un operador
III) ***AT+CREG:*** Registrarse en una red
IV) ***AT+WOPN:*** Leer nombre del operador

3) Comandos de seguridad:
I) ***AT+CPIN:*** Introducir el PIN
II) ***AT+CPINC:*** Obtener el número de reintentos que quedan
III) ***AT+CPWD:*** Cambiar password

4) Comandos para la agenda de teléfonos
I) ***AT+CPBR:*** Leer todas las entradas
II) ***AT+CPBF:*** Encontrar una entrada
III) ***AT+CPBW:*** Almacenar una entrada
IV) ***AT+CPBS:*** Buscar una entrada

5) Comandos para SMS
I) ***AT+CPMS:*** Seleccionar lugar de almacenamiento de los SMS
II) ***AT+CMGF:*** Seleccionar formato de los mensajes SMS:
- *Modo texto*
- *Modo PDU*

III) ***AT+CMGR:*** Leer un mensaje SMS almacenado
IV) ***AT+CMGL:*** Listar los mensajes almacenados
V) ***AT+CMGS:*** Enviar mensaje SMS
VI) ***AT+CMGW:*** Almacenar mensaje en memoria
VII) ***AT+CMSS:*** Enviar mensaje almacenado
VIII) ***AT+CSCA:*** Establecer el Centro de mensajes a usar
IX) ***AT+ WMSC:*** Modificar el estado de un mensaje

sea de considerables costos si nuestra solución de negocios debe manejar cientos de miles de mensajes por hora.
Adicionalmente requiere la generación de contratos con cada una de las TELCOs mediante las cuales deseemos ofrecer la solución a los clientes finales.
En este caso, también deberíamos desarrollar un framework que reciba y envíe mensajes SMS mediante SMPP o utilizar alguno de los que existen en el mercado.

Conclusión de arquitecturas SMS
Existen por supuesto, arquitecturas adicionales y aplicaciones novedosas que requieren de integraciones ingeniosas para llevarlas a cabo mediante los protocolos que intervienen en el envío y recepción de SMS. Quedará al análisis de cada requerimiento, la decisión de la arquitectura final. Adicionalmente, hay que tener en cuenta antes de ponerse a escribir código, de que probablemente ya exista un framework SMS que puede resolvernos la implementación técnica de nuestro negocio.

Soluciones WAP/GPRS remotas con un servidor Web

Casi olvidadas por el usuario final, y adoptadas de una forma muy light por parte del mercado corporativo, se encuentran aquellas aplicaciones que permiten "navegar" en el teléfono por páginas Web remotas que pueden brindar las mismas o más soluciones que los mensajes SMS.
Las aplicaciones remotas que utilizan transferencia de datos por WAP/GPRS, tienen ventajas gráficas importantes y una versatilidad que permite que sean estándares en casi la totalidad de los dispositivos móviles.

Ejemplos de aplicaciones podrían ser:
- La implementación en versión móvil de cualquier CRM, ERP o sistema de gestión utilizado por las empresas.
- Sistemas de relevamiento y control.
- Sistemas de búsqueda de servicios.
- Y en general, cualquier aplicación, de uso general o privada, que implementada en un sistema Web estándar, permita ser migrada a las pantallas reducidas de los celulares.

Para la implementación de soluciones remotas, es necesario contar con un servidor Web como por ejemplo Microsoft IIS y desarrollar una aplicación Web para dispositivos celulares. Con Microsoft .NET 2005 encontraremos numerosas ventajas en la generación de este tipo de aplicaciones.
La ventaja radica en la facilidad de desarrollo y en la simplicidad de la comunicación entre el dispositivo móvil y nuestro servidor de aplicaciones Web. Obteniendo ventajas en el tiempo, costo y versatilidad de los sistemas implementados.

Lectura Adicional

- www.behindsystems.com/frameworks/mobile/mobilefwk.htm
- msdn.microsoft.com/mobility/default.aspx
- java.sun.com/javame/index.jsp



Otras Arquitecturas

Existen por supuesto, diferentes posibilidades a la hora de implementar una solución para telefonía móvil. Algunas de ellas son:

Aplicaciones Java locales en el celular
Consiste en realizar una aplicación local que debe instalarse en el celular, la cual será implementada con J2ME. Gran cantidad de juegos para celulares están implementados en esta plataforma. Podría también implementarse soluciones de negocio, obteniendo como ventaja la capacidad gráfica del motor, y el procesamiento local en el dispositivo.
Por supuesto, la aplicación podría utilizar WAP/GPRS, SMS, Bluetooth u otros canales para luego comunicarse con el servidor de aplicaciones si así fuese necesario.
Como desventaja puede citarse la necesidad de testear la aplicación en cada modelo de celular, debido a que actualmente es difícil encontrar, a menos que se utilice un reducido framework estándar, compatibilidad entre muchos modelos.

Aplicaciones .NET locales en el celular
La alternativa Microsoft a las aplicaciones locales de Java, es el Compact .NET Framework. En este caso, el dispositivo debe ser inteligente y tener el sistema operativo Windows CE o Windows Mobile. Encontraremos que utilizando esta plataforma, podremos simular aplicaciones Windows de escritorio en el dispositivo móvil, obteniendo ventajas de rendimiento, seguridad, acceso a datos y facilidad de desarrollo, comunicación y mantenimiento.
Con Microsoft .NET 2005 es muy accesible la creación de este tipo de aplicaciones, que en general serán utilizadas de forma corporativa.

Aplicaciones embebidas en el SIM del celular: GSM Toolkits
Por último, existe la posibilidad de agregar aplicaciones embebidas en el SIM del celular mediante GSM Toolkits.
Aquí estamos hablando de sistemas core que sirven de frameworks para resolver arquitecturas comunes a muchas aplicaciones y que, por motivos de negocio, son almacenados en los SIM para que estén disponibles en todos los móviles.
Dejaremos un análisis profundo de este tema por demás interesante para futuras entregas.

Conclusión

El desarrollo de aplicaciones para dispositivos móviles está recién empezando. No se ha alcanzado a obtener el máximo rendmiento de las arquitecturas y protocolos actuales, y surgen por año varios nuevos estándares que agregarán mayores prestaciones a los celulares.
En esta nota, inicié el camino hacia un conjunto de términos y soluciones, que brindan posibilidades excepcionales a nuestros requerimientos de negocio y comunicación en general. ●



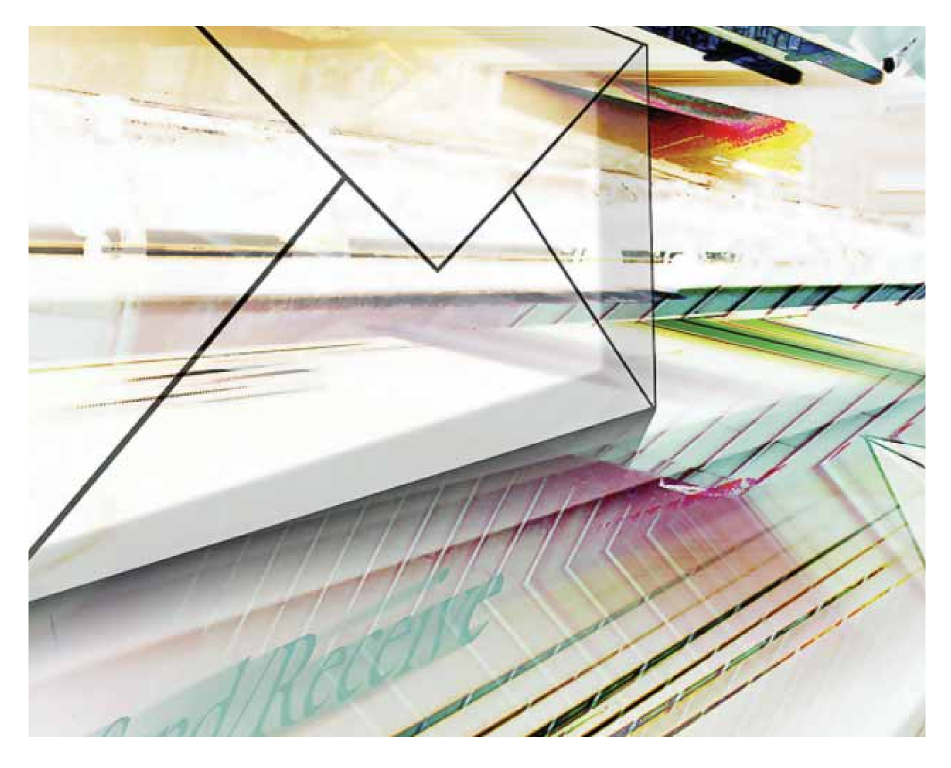
Para más información comuníquese con el Centro de Atención al Cliente (011) 6770-7669 de Lunes a Viernes de 9:00 hs. a 18:00 hs. o por e-mail: ait@ar.sony.com.

Administración de Correo NO deseado

Por: **Jim McBee** MCSE, MCT y Microsoft Exchange MVP

Las actividades diarias de un administrador de mensajes incluyen muchas tareas tanto pequeñas como grandes, incluyendo la administración del buzón y de los destinatarios, administración de colas de espera, protección contra virus y supervisión del sistema. Si todo esto no fuera suficiente, las personas que envían spam empeoran aún más las cosas, pero hay formas de solucionarlo.

En muchas organizaciones la prevención de correo no deseado y la administración de cuarentena pueden tomar la mayor parte del día de un administrador; algunas organizaciones tienen empleados de tiempo completo dedicados a la administración de las tareas relacionadas con los spam. En organizaciones con pocas tecnologías de protección contra correos no deseados, los cálculos indican que cada uno de los empleados puede invertir hasta 45 minutos al día borrando los spam de sus buzones. Uno de los temas más populares en las conferencias, en grupos de noticias de Internet y en las juntas es cómo prevenir el correo no deseado.



Hay muchos productos y enfoques que se pueden utilizar al intentar eliminar el spam. Cuando busca los enfoques que se van a combatir, una organización cuenta con varias opciones, incluyendo:

- Utilizar software del lado del cliente.
- Utilizar software en el servidor de correo o que se integra con el software antivirus del servidor de correo.
- Implementar una puerta de enlace contra correo no deseado que se localice en la red de perímetro de la organización.
- Solicitar que un tercero realice las funciones iniciales de higiene de mensajes (software contra correo no deseado y antivirus) antes de que el mensaje sea enviado a los servidores de correo de la organización

En este artículo exploraremos brevemente la experiencia de una organización para encontrar la solución contra correo no deseado indicada para satisfacer sus necesidades de negocios.

¡Ayuda! ¡Estamos inundados!

La Compañía X es una pequeña compañía de consultoría; Microsoft Exchange Server 2003 brinda soporte a sus 18 buzones. El correo electrónico es la parte vital de esta compañía.

Por algunos años la Compañía X no contaba con un especialista en informática e ignoró por completo el problema cada vez mayor que representaba el correo no deseado. Con frecuencia, cada uno de los empleados instalaba y configuraba sus propias soluciones contra correo no deseados del lado del cliente. Un empleado reportó que en un solo período de 24 horas recibió 750 mensajes que él clasificó como basura.

¿Pero qué tan grave era el problema? Cuando la compañía actualizó su servidor a Exchange Server 2003 Service Pack 2, utilizó el Filtro Inteligente de Mensajes (IMF) integrado, los contadores del monitor del sistema de IMF y los contadores del monitor del sistema SMTP para cuantificar mejor qué tan grande era el problema del spam. La organización estableció el umbral de la puerta de enlace de IMF a un nivel de confianza de correo no deseado (SCL) de 5 y lo configuró para su archivo. Durante 24 horas, el monitor del sistema registró las estadísticas de IMF, éstas se muestran en la Figura 1.

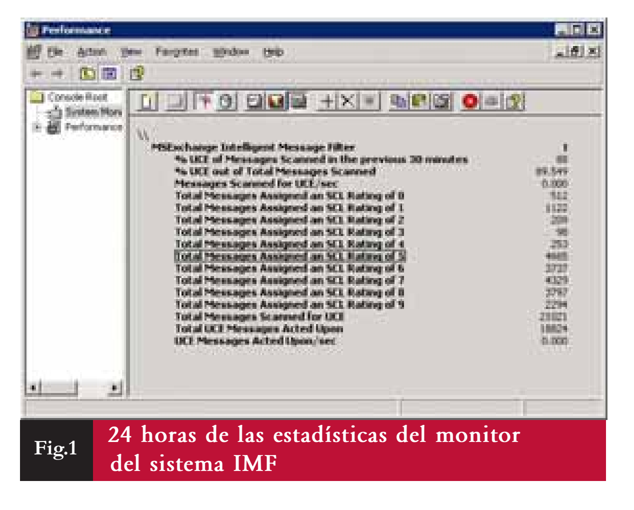


Fig.1 24 horas de las estadísticas del monitor del sistema IMF

En 24 horas este servidor de correos aceptó 21.021 mensajes, de los cuales, 89.5 por ciento tenían un SCL de 5 o superior. Este porcentaje de mensajes consumieron cerca de 35 megabytes (MB) de espacio en disco en el archivo, y se requirió un aproximado de 40 MB de ancho de banda de Internet para recibirlos. Ésta es realmente una cantidad alarmante de spam para una compañía tan pequeña.

Se implementaron algunas medidas iniciales para intentar combatir la afluencia del spam, incluyendo la habilitación de listas de bloqueo en tiempo real (RBLs) y filtros de destinatarios. Durante un período de 5 días se supervisaron las estadísticas del Receptor de filtro de transporte para determinar qué tan efectivos eran los proveedores de RBL y cuántos mensajes se rechazaban debido a que eran enviados a destinatarios inválidos. La Figura 2 muestra las estadísticas de la Pileta del filtro de transporte, de aproximadamente 211 mil consultas de la lista de bloqueo DNS, se rechazaron 111 mil conexiones debido a que la dirección IP del remitente se encontraba en una de las listas de bloqueo DNS.

Se rechazaron 76 mil mensajes adicionales debido al filtro de destinatarios, ya que los destinatarios del mensaje no eran válidos en el servicio de Active Directory. Claramente, alguien estaba intentando entregar una gran cantidad de mensajes a destinatarios inválidos. Un simple vistazo a las Sesiones actuales del Servidor virtual de Exchange SMTP (Figura 3) mostró docenas de sesiones SMTP de entrada en un momento dado.

Se revisaron los registros del protocolo SMTP y se determinó que casi todas estas sesiones SMTP estaban siendo enviadas a destinatarios inválidos. Este tipo de envío de spam se conoce como "envío de correo no deseado de diccionario" debido a que el software de la persona que envía los spam utiliza un diccionario de nombres comunes. En el cuadro se muestra un ejemplo del registro.

De aproximadamente mil entregas que se intentaron y que fueron supervisadas desde esta sesión en este archivo de registro, sólo uno era un usuario válido. Este usuario fue probablemente el "caso de índice" y es como la persona que envía los correos no deseados descubrió el dominio de la compañía en primer lugar.

A pesar de las medidas de protección iniciales, tales como utilizar listas de bloqueo en tiempo real, una trampa de brea (tar pit) SMTP (que inserta un retraso en las comunicaciones de SMTP asociadas con el spam), y el Filtro Inteligente de Mensajes ayudaron a reducir en gran medida la cantidad de correos no deseados que recibían los usuarios, el hecho siguió siendo que gran parte de los correos no deseados siguieron filtrándose al servidor Exchange donde se ponían en cuarentena o pasaban a la carpeta principal de Correo electrónico basura del usuario. El spam siguió consumiendo recursos del servidor, la red y el usuario.

\\	
MSExchangeTransport Filter Sink	1
Block List DNS Queries Issued	211654
Connections Rejected by Block List Providers	111393
Connections Rejected by Deny List	211
Messages Filtered by Sender Filtering	463
Messages Filtered Due to Blank Sender	15
Recipients in Connection Filtering Exception List	49
Recipients Rejected by Directory Lookups	76044

Fig. 2 Estadísticas del transporte de mensajes de SMTP

```
21:15:14 212.183.55.210 telekom.at EHLO - +telekom.at 250
21:15:15 212.183.55.210 telekom.at MAIL - +From:+<vztt@telekom.at> 250
21:15:53 212.183.55.210 telekom.at RCPT - +To:<arring@somorita.com> 550
21:16:08 212.183.55.210 telekom.at RCPT -  +To:<capps@somorita.com> 550
21:16:24 212.183.55.210 telekom.at RCPT - +To:<carl@somorita.com> 550
21:16:39 212.183.55.210 telekom.at RCPT - +To:<carlis@somorita.com> 550
21:16:54 212.183.55.210 telekom.at RCPT - +To:<blear@somorita.com> 550
21:17:10 212.183.55.210 telekom.at RCPT - +To:<brewer@somorita.com> 550
21:17:25 212.183.55.210 telekom.at RCPT - +To:<boykin@somorita.com> 550
21:17:39 212.183.55.210 telekom.at RCPT - +To:<burris@somorita.com> 550
21:17:55 212.183.55.210 telekom.at RCPT - +To:<carl@somorita.com> 550
21:18:12 212.183.55.210 telekom.at RCPT - +To:<butts@somorita.com> 550
21:18:27 212.183.55.210 telekom.at RCPT - +To:<bill@somorita.com> 550
21:18:41 212.183.55.210 telekom.at RCPT - +To:<blue@somorita.com> 550
21:18:57 212.183.55.210 telekom.at RCPT - +To:<bower@somorita.com> 550
21:19:12 212.183.55.210 telekom.at RCPT - +To:<castle@somorita.com> 550
21:19:27 212.183.55.210 telekom.at RCPT - +To:<atwood@somorita.com> 550
21:19:42 212.183.55.210 telekom.at RCPT - +To:<ash@somorita.com> 550
```

Ejemplo de Registro

Evaluar las soluciones

Cuando la Compañía X estaba evaluando las soluciones, emplearon diversos criterios a través de los cuales eligieron una solución. Estos criterios incluyeron:

• Reducir la cantidad de spam que deben ser procesados por el servidor Exchange.

• Reducir la cantidad de spam que entran al buzón del usuario de manera que no se sincronicen con los dispositivos Windows Mobile, no se descarguen a los buzones del usuario remoto y no provoquen que se genere un mensaje Fuera de la oficina.

• Utilizar una solución que se pueda actualizar con frecuencia de manera que tanto las nuevas técnicas de spam como las amenazas de virus del día cero se puedan detectar.

• Proteger el servidor Exchange de manera que no esté directamente expuesto a Internet.

• Asegurarse de que la solución sea escalable y tolerante a fallas para que el correo no se pierda durante el proceso de inspección y entrega.

• Reducir la cantidad de ancho de banda de la red y los recursos del sistema que consumen los spam.

• Requerir el mínimo esfuerzo administrativo debido a que la Compañía X no cuenta con un personal de informática dedicado.

La solución que eligió el Cliente X fue utilizar un proveedor administrado. Así redirige sus registros DNS MX públicos a los centros de datos del proveedor administrado, el proveedor ofrece servicios de higiene de mensajes tales como protección contra spam y antivirus, y el proveedor envía posteriormente los mensajes al servidor Exchange de la Compañía X.

La Compañía X puede elegir su tolerancia en cuanto a qué tan agresivamente filtran el correo no deseado potencial los sistemas del proveedor administrado. Muchos usuarios que veían cientos de mensajes de correo no deseado en sus Bandejas de entrada todos los días, ahora reportan que por lo general reciben cinco o menos mensajes de spam al día. ●

La Compañía X se basa en una empresa pequeña real que se vio inundada con correo no deseado.

Este artículo ha sido publicado previamente en el TechNet de ***Microsoft****: http://technet.microsoft.com/*

Current Sessions

User	From	Connected Time
alicedsl.de	222.137.137.56	3349 seconds
angelfire.com	80.171.1.10	93 seconds
	221.212.249.187	61 seconds
charlottenc.nefn.com	80.34.39.237	42 seconds
furmanselz.com	210.221.57.10	41 seconds
sbcglobal.net	220.75.52.226	40 seconds
cox.net	88.7.169.173	37 seconds
lycos.com	59.16.174.56	37 seconds
angelfire.com	221.163.128.184	36 seconds
mojozone.co.nz	65.245.103.77	35 seconds
moen.com	220.73.2.40	32 seconds
biddeford.com	58.234.95.25	32 seconds
finh.com	68.148.211.72	31 seconds
mitra.com	222.121.145.195	30 seconds
maine.rr.com	218.64.70.226	30 seconds
visualink.com	220.169.56.69	24 seconds
alltel.net	219.250.200.136	22 seconds
netvip.com	222.121.35.130	21 seconds
bellsouth.net	200.104.245.14	18 seconds
execpc.com	82.38.89.202	16 seconds
mindspring.com	222.115.168.103	15 seconds
higherreturnmedia.com	64.239.70.32	15 seconds
winn-cache-4.server.n...	82.26.30.138	14 seconds
lightlink.com	221.150.82.44	13 seconds
	69.142.58.121	13 seconds
ix.netcom.com	219.250.200.136	13 seconds
knittingbasket.com	200.77.198.47	10 seconds
artware.com	220.81.233.150	10 seconds
modeteam-lorette.de	66.108.230.193	6 seconds
rrm.co.uk	84.165.125.98	3 seconds
adelphia.com	147.47.247.61	3 seconds
bellsouth.net	69.141.164.181	3 seconds

Fig. 3 Sesiones SMTP de entrada

RFID Radio Frecuency Identification



En la lucha eterna del equilibrio entre la seguridad y la funcionalidad, hemos visto pasar a varias tecnologías, pero como no podía ser de otra manera le llegó el turno a RFID (Radio Frequency Identification).

RFID es una tecnología de identificación por radiofrecuencias que permite el reconocimiento automático a distancia, basado en uno de sus principales componentes, los TAGS (Etiquetas) de RFID, permitiendo así un beneficio muy importante en lo que refiere a la logística, la distribución y la cadena de abastecimiento de las organizaciones, pero como veremos más adelante, la aplicación de esta tecnología también esta siendo adoptada en muchos otros aspectos y procesos, como el control de accesos y el pago electrónico y la identificacion de documentación personal.
Un Sistema de RFID suele basarse en varios componentes: Tags, Tag Readers, Front-Ends, Middleware, Back-Ends.

La Tecnología y sus componentes

Esta tecnología permite la transmisión de información a distancia gracias a la etiqueta RFID (TAG), la cual al ser leída por el Lector de RFID transmite la información contenida en ella a la aplicación intermedia (Middleware). Ésta se encarga de procesarla para, finalmente, tomar o depositar la información en una base de datos, típicamente ubicada en el back-end (figura 1). Esa información transmitida por el Tag puede proveer información relacionada con la identificación del producto, la ubicación de la misma o bien otros datos específicos, tales como color, precio, datos de compra, datos de vencimientos entre otros.

Tipos de Etiquetas

Existen distintos tipos de etiquetas (figura 2) y se diferencian entre sí por la frecuencia en la que operan, la cantidad de información que pueden contener, el tipo de funcionamiento y su durabilidad. Existen tres tipos:

Otro round entre la funcionalidad y la seguridad

RFID es una tecnología de identificación por radiofrecuencia, la cual si bien esta presente desde un largo tiempo atrás, es hoy que las empresas ven en ella un actor principal en cuestiones como la logística y la cadena de abastecimiento, pero también es cierto, que desde la óptica de la Seguridad también son visibles varios de los riesgos a tener en cuenta en su utilización.

Por: **Ezequiel Martin Sallis** - Senior Security Consultant - **CISSP Certified**

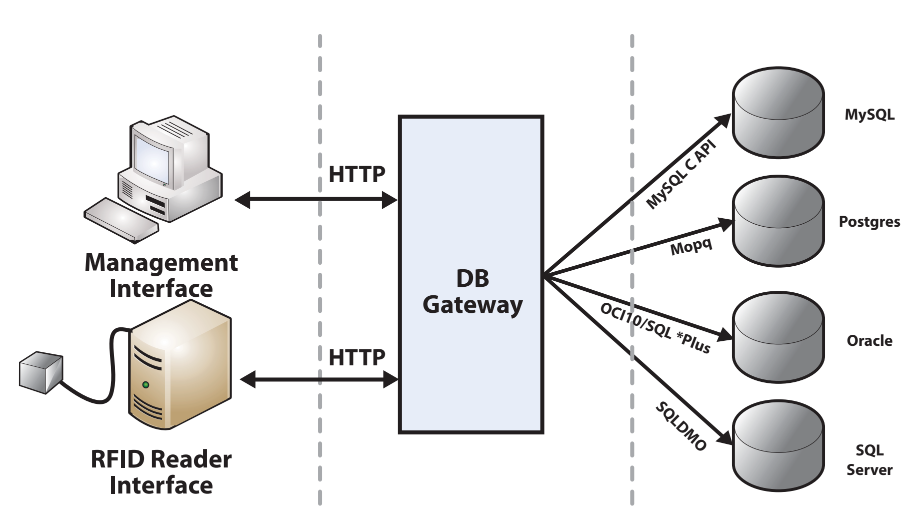


Fig.1 Arquitectura Basica RFID

Etiquetas Pasivas:
Estas operan en la Frecuencia de los 13,56 MHZ y no tienen fuente de energía interna, sino que la pequeña corriente inducida en la antena, brindada por la señal entrante de la frecuencia radial, produce la energía suficiente para que el circuito integrado pueda encenderse y comenzar a transmitir (Backscatter).
Estas etiquetas son las de menos tamaño, por ende las más livianas y con una vida útil que puede ir hasta los 99 años.

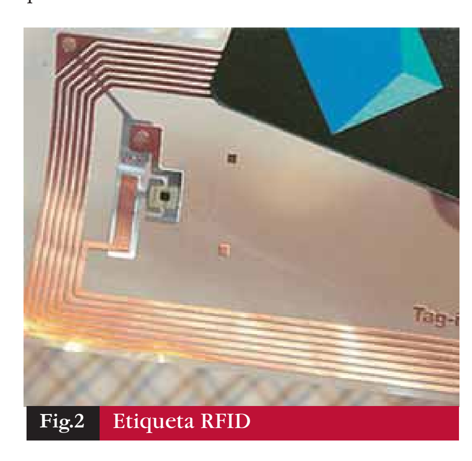


Fig.2 Etiqueta RFID

Etiquetas Semipasivas:
Son muy similares a las Pasivas, salvo por el agregado de una pequeña batería. Esta batería mantiene una corriente continua en la memoria no volátil del circuito integrado, por lo cual la antena no debe preocuparse por recolectar dicha corriente. La antena esta más optimizada a su función de transmisión de radio frecuencia lo cual hace que sea más rápida y robusta que los Tags Pasivos.

Etiquetas Activas:
Las etiquetas activas poseen su propia fuente de energía y son capaces de alcanzar mayores distancias (10 metros aproximadamente). Además, al poseer una batería cuya vida útil de es de hasta 10 años, estos economizan el consumo de energía, trabajando en intervalos definidos de operación.

Tipos de Frecuencias

Existen distintas frecuencias en las que los sistemas de RFID pueden operar. Cada una de ellas representa distintos pro y contras en base a su aplicación.

Low Frequency (125 a 134.2 kHz y de 140 a 148.5 kHz)
Las etiquetas y lectores de baja frecuencia se encuentran típicamente en tarjetas utilizadas para el control de acceso (Contact less Smartcards). La distancia en este caso es muy acotada y esta limitada a centímetros.

High Frequency (13.56 MHz)
Esta frecuencia opera en distancias de hasta un metro y se utiliza típicamente en la identificación de productos o personas (pacientes, convictos y otros).

Ultra-High Frequency (915 MHz, 433.92 MHz. o 315 MHz)
Dependiendo la tecnología pueden llegar a operar en una distancia de hasta 10 metros o más, típicamente esta tecnología es la que se utiliza para las cadenas de distribución y abastecimiento.

Microwaves
Utilizadas para grandes distancias y mayor velocidad, operan en el rango que va de los 30 metros a los 100 metros, un lugar donde se la utiliza suele ser por ejemplo los sistemas de pase automático de las autopistas.

Aplicaciones RFID

Hoy en día existen numerosas aplicaciones para estas tecnologías (Ver figura 3 y 4), pero la más creciente es la que esta bajo el estándar EPC (Electronic Product Code), utilizada en la identificación de productos, la cual brinda una clave única para un producto o Pallet, que permite detallar información sobre el mismo en cualquier momento de la cadena de abastecimiento.
Adicionalmente, entre otras aplicaciones podemos mencionar las siguientes:

- Implementaciones ganaderas para la identificación de ganado, su historial, sus progenitores, sus descendientes y su producción.

• Identificación en medicamentos de la fecha de vencimiento o bien la información sobre los efectos secundarios del mismo.
• Medios de pago electrónico (Mastercard Paypass).
• Identificación de pacientes.
• Identificación de convictos.
• Identificación de billetes de alta denominación.
• Identificación de pasaportes.
• Identificación de registros de conducir.
• Identificación de entradas a eventos deportivos y espectáculos (Mundial Alemania 2006).
• Sistemas de Control de acceso.
• Otras, muchas otras aplicaciones...

Podemos agregar que ya existen implementaciones de RFID mandatorias, por ejemplo, entre algunas de las empresas y organizaciones que han emitido su mandato de implementación podemos mencionar al DOD (Departamento de Defensa de los Estados Unidos) y a Wal-Mart. Este último, obliga a todos sus proveedores a colocar los Tags de RFID en todos los productos que tengan como destino final la góndola de Wal-Mart, impactando de esta manera en miles de compañías alrededor de todo el mundo. La fecha límite fue postergada en varias ocasiones, debido a que muchos vendedores tuvieron dificultades al implementar los sistemas de RFID.

Riesgos - Desde la invasión a la privacidad hasta SQL Injection

Tal como mencionamos en nuestra introducción, nos encontramos nuevamente en la dificultad de buscar el equilibrio entre la funcionalidad de esta tecnología y los riesgos que ésta puede introducir desde la óptica de la seguridad de la información.
Es por esto que a continuación haremos un breve analisis de los riesgos desde dos ópticas bien diferenciadas, por un lado, el tema de la privacidad vs. RFID y por otro, desde un punto de vista bien técnico algunas de las técnicas de ataques ya presentes contra algunas implementaciones de esta tecnología.

Privacidad

Los especialistas piensan en cómo transformar esta tecnología en la herramienta para poder establecer y entender el perfil del consumidor tan buscado y tratar de personalizar sus productos, sus mensajes y sus descuentos para acrecentar sus ventas. Es por esto que ya, tanto en Estados Unidos como así también en algunos lugares de Europa, se han levantado movimientos en contra de esta tecnología, argumentando que la misma invade la privacidad de los ciudadanos. A decir verdad, este tipo de cosas nos llevarían a preguntarnos por ejemplo:
• Y si comprase medicamentos RFID-tagged, como por ejemplo anti-depresivos, ¿quisiera que alguien que pase caminando a mi lado (o no tan a mi lado) lo sepa?
• Y si estoy en mi casa y alguien pasa con su auto y un lector de RFID, ¿puede determinar todo lo que he comprado hasta el momento y establecer mis características de consumidor?
• Y con esta tecnología en mi ropa, ¿alguien puede determinar con precisión donde encontrarme?
Igualmente, y con sinceridad, si nosotros continuamente compramos con tarjetas de créditos, utilizamos tarjetas shopping para los descuentos, damos nuestros datos a cambio de una remera de Merchandising de una compañía, permitimos a las páginas Web setear cookies en nuestras PCs... y siendo que todo esto permite potencialmente realizar un seguimiento sobre nosotros, hacer un estudio de nuestros consumos, poder ubicarnos en un determinado momento... ¿De que nos preocupamos entonces con el RFID? En fin, si bien quizás este conflicto se demore un tiempo en llegar por estos lados, en el país del norte ya existe gran cantidad de legislación encargada de proteger la privacidad de los consumidores y otros aspectos relacionados con la utilización de la tecnologia RFID.
Entre algunas de las leyes existentes podemos mencionar las siguientes, a modo de ilustrar un poco más la relevancia que se le da a esta problemática:



Fig.3 Mastercard Paypass

California - SB1834
Restringe la manera en que los comercios de California utilizan los Tags de RFID, en pos de que los tags de sus productos no sean utilizados para la identificación de un individuo. Junio 25, 2005.

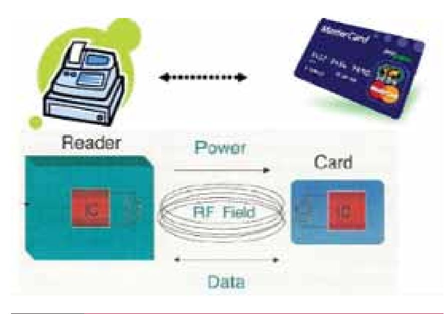


Fig.4 Comunicación Smartcards contactless

Massachussets – HB 1447, SB 181
Requiere de la advertencia al usuario sobre la existencia de etiquetas de RFID en los productos que adquiere, como así también, indicar el procedimiento para realizar la remoción del mismo. Por otro lado limita la información de la etiqueta a aspectos de inventario solamente.

New Hampshire – HB 203
Requiere la comunicación escrita o verbal, por parte del comercio de que el producto que vende contiene una etiqueta.

Rhode Island – HB 5929
Prohíbe la utilización de RFID para la identificación y el seguimiento de estudiantes, empleados y clientes con fines que beneficien al negocio.

Utah – HB 185
Enmienda la ley de delitos informáticos para la inclusión de la tecnología RFID.

Riesgos técnicos

Los sistemas RFID se relacionan con varios procesos críticos, como el control de acceso físico, el seguimiento de productos, los sistemas de pago y otros más.
Si bien, como vimos, los riesgos más publicitados de esta tecnología se relacionan con la privacidad, a continuación veremos algunos otros que también deberían ser considerados.

• Relay Attacks en tarjetas de proximidad.
• Destrucción del TAG.
• Prevención de Lectura.
• RFID - SQL Injection.
• Spoofing.

Relay Attacks en Tarjetas de Proximidad

Un Sistema de Smartcard sin contacto (ISO 14443A) se comunica con otros dispositivos sin la necesidad de contacto físico. Lo hace a través de RFID; la tarjeta es pasiva y es activada por el lector para realizar la transferencia de datos.
Las implementaciones de estos sistemas en pasaportes, medios de pago electrónico y tarjetas de acceso (típicamente Físico) están incrementándose día a día. Uno de los puntos más importantes en relación a esta tecnología y a la seguridad es, por un lado, su corto rango de operación (10 centímetros aproximadamente), y por otro, el hecho de que la comunicación entre un extremo y el otro va cifrada (Ver Figura 4).
Lo que se ha logrado en un estudio llevado a cabo por la Universidad de Tel Aviv (Israel), es que creando un lector falso (Leech) y una tarjeta falsa (Ghost) uno podría impersonar un usuario o una transacción, prácticamente sin límites de distancia.
El concepto de este ataque nace por un ataque

basado en memory cards, las cuales no tienen la capacidad de procesar la información y la almacenan en una banda magnética, que es leída y procesada por el lector, mediante un **POST Interceptor**. Y, gracias a la utilización de algoritmos de encriptación débiles **"Static Data Authentication" (SDA)**, la información de la tarjeta y el pin era almacenado en el POST Interceptor para una posterior utilización fraudulenta.
Es en base a lo anterior, a la creciente implementación de smart cards y a sus variados usos que los atacantes optan por el **Ataque de Relay** y utilizan la información en **Real Time**.
Pongamos un ejemplo, un usuario de una organización que cuenta con una smartcard contactless para el acceso físico a un área restringida, típicamente podría llevar su tarjeta colgando en la cintura como habitualmente se hace. Éste podría ser sorprendido por un potencial atacante en un recinto donde exista poco espacio y gran cantidad de gente (ascensor, subterráneo). Allí el potencial atacante podría acercarse lo suficiente como para lograr activar el smartcard del usuario mediante un Smartphone. Modificado para actuar y emular a un lector (Leech), una vez logrado esto, el atacante transmitiría en tiempo real vía GRPS, los datos a otros dispositivo similar, en este caso, la tarjeta falsa (Ghost), la cual por último y estando cerca del lector original, transmitirá los datos recibidos y brindará acceso al intruso al área restringida (Ver Figura 5). Es importante destacar que en la mayoría de este tipo de ataques la distancia no importa, como así tampoco el cifrado de datos, ya que en la mayoría de los casos éste no detecta el MITM (Man in The Middle).

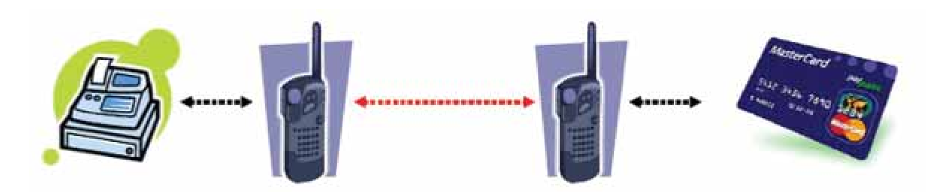


Fig.5 Smartcard Relay Attack

Destrucción del TAG

Este tipo de ataque requiere poco conocimiento técnico para ser realizado. La idea es destruir el tag pasivo colocado en un producto, de manera que éste no pueda ser identificado ni tampoco se permita realizar su seguimiento. Típicamente este tipo de acciones esta motivada en defensa de la privacidad de los consumidores.
Técnicamente, ésto podría llevarse adelante cortando la antena o bien friendo el tag en un microondas, pero de seguro que el producto también resultaría destruido, es por eso que se creo el "**RFID-Zapper**", el cual utiliza el método del microondas pero en una dosis menor, evitando así la destrucción del producto mediante la generación de un campo electromagnético fuerte, que desactivara el chip para siempre.

Prevención de Lectura

La idea de esto no es desactivar el tag pasivo, sino prevenir su lectura. El fin de esta práctica es la protección de la privacidad, o bien, como hemos visto anteriormente también podrían prevenir ataques de Relay contra las tarjetas de proximidad (imagínense cuán cerca podría estar alguien de su tarjeta en un ascensor).
Existen dos materiales que impiden la lectura de RFID: el metal y el agua, pero mucho más cómodo es darle un nuevo uso al famoso **SILVER TAPE**.

Spoofing

Un atacante podría escribir en un TAG en blanco datos válidos que impersonen a un producto. De hecho, ya se realizaron dos ataques (en ambiente controlado): uno se basó en sniffear, descifrar y spoofear un dispositivo RFID utilizado para cargar gasolina, y el otro se utilizó para desactivar un sistema de alarmas para automóviles basado en la inmovilización del mismo.

SQL Injection

Si bien se trata de un ataque ya conocido y muy utilizado hoy en día, esta nueva tecnología no queda fuera del alcance del mismo, este ataque se basa en la inyección a la base de datos (Back-End) de sentencias SQL especialmente creadas, que podrían permitir acceder a información no autorizada, realizar modificaciones o borrar una tabla, entre otras cosas.
En base a la estructura de una implementación RFID (Figura 1), esto es posible por falta de controles de seguridad adecuados en el middleware y en la base de datos. En este tipo de ataques el potencial intruso debería grabar previamente en un tag en blanco, del tipo activo, una sentencia de SQL previamente armada, con el fin de que cuando este TAG sea leído por el lector, éste pase los datos al middleware y posteriormente esta sentencia se ejecute en la base de datos. Otro aspecto a tener en cuenta es la cantidad de datos que pueden entrar en un tag. Esto no es un limitante ya que con pocas cantidades de información se puede causar un gran impacto.
De la misma manera que se inyecta esta sentencia, se han realizado en ambiente de laboratorios pruebas en las cuales se logró con éxito crear y replicar el primer worm que utiliza la tecnología RFID.
Por último, cabe aclarar que este tipo de ataques no tiene relación directa con la tecnología RFID en sí misma, sino que al igual que en las aplicaciones Web y similares, la falta de validación de entrada de datos y la falta de metodologías seguras de programación hacen que esto sea posible.

Conclusión

Como hemos podido leer a lo largo de este artículo, la incorporación de nuevas tecnologías en lo procesos del negocio, no solo representan un análisis funcional y económico, sino también que la seguridad debe ser uno de los puntos a tener muy en cuenta.
La creatividad que tienen este tipo de ataques hacen que los controles de seguridad existentes puedan no tener sentido. Es por eso que como profesionales de seguridad tenemos que desarrollar también el sentido de la percepción y de la creatividad. ●

Sobre el Autor: Ezequiel Sallis

Ezequiel Sallis es miembro de ISSA (Information System Security Association), como así también el miembro fundador del capítulo de ISSA en Argentina www.issaarba.org . Es un profesional certificado CISSP, con varios años de experiencia de campo. Actualmente se desempeña en el cargo de Director de Research & Development en I-Sec Information Security Inc.

Referencias y Lecturas Complementarias

- *Fundamentals and Applications in Contactless Smartcards & Identification*
  Klaus Finkenzeller
- *Python library for exploring RFID devices*
  http://rfidiot.org
- *Practical Relay Attacks Against ISO 14443 Proximity Cards*
  Gerhard Hancke & Dr Markus Kuhn
- *Low Cost Attacks on Tamper Resistant Devices*
  Ross Anderson & Markus Kuhn
- *A New Approach to Hardware Security Analysis in Semiconductors*
  Sergi Skorobogatov
- *RFID Essentials*
  O'Reilly
- *Texas Instruments DST attack*
  www.jhu.edu/news_info/news/home05/jan05/rfid.html
- *RFID relay attacks*
  www.cl.cam.ac.uk/~gh275/relay.pdf
- *RFID virus*
  www.rfidvirus.org/papers/percom.06.pdf
- *Smartdust*
  http://en.wikipedia.org/wiki/smartdust

Ethernet in the First Mile

Miguel F. Lattanzi
Ing. en Telecomunicaciones (IUPFA)
SIEMENS – Red de Acceso y Datos

EFM representa el esfuerzo realizado para llevar la tecnología Ethernet desde el límite del abonado hasta la red del Proveedor, permitiéndo aprovechar las nuevas aplicaciones multimedia a costos más accesibles.

La Primera Milla

Ethernet in the First Mile, o EFM, es un estándar (IEEE 802.3ah) que define los protocolos, el modelo de capas y las topologías a utilizar en las redes de acceso, que implementen Ethernet como tecnología de transmisión de datos.
La primera milla o red de acceso es el conjunto de enlaces/equipos que se encuentran entre los equipos de los abonados y los equipos de la red pública del proveedor de servicios, ya sea en una CO (Central Office) o en un PoP (Point of Presence).
Estos enlaces siempre han sido un inconveniente, en cuanto a velocidad de transmisión, dado que tanto las redes LAN (Local Area Network) de los usuarios, así como el core –núcleo– de la red del proveedor, siempre tuvieron una mayor velocidad de transmisión de datos. Debido a esta diferencia de velocidades, se genera en la red de acceso el efecto conocido como "cuello de botella", afectando a todo el tráfico que transita por ella. Otro inconveniente que se presenta en la primera milla es la conversión de protocolos que se debe realizar, dado que los protocolos usados por los usuarios en la LAN y los utilizados en la red de acceso son distintos. En esta última se encuentran escenarios basados en xDSL (x Digital Subscriber Line), ISDN (Integrated Services Digital Network) y cable-modem, entre otros. La conversión de protocolos requiere de equipos extra que elevan los costos, demanda tiempo de procesamiento e impacta directamente sobre el diseño de las redes, haciéndolo más complicado.
EFM pretende eliminar todos estos problemas, al mismo tiempo que genera otras ventajas extras, entre ellas podemos mencionar el costo de los equipos Ethernet, que es muy inferior al de otras tecnologías.

Modelos de Conexión

EFM plantea tres escenarios posibles de conexión reutilizando los enlaces de cobre existentes en la mayoría de las redes de acceso, y usando los enlaces de fibra óptica instalados más recientemente.
Una de las posibilidades que plantea IEEE 802.3ah, es Ethernet punto a punto sobre enlaces de cobre, que da lugar a EoVDSL (Ethernet over Very-high-bit-rate Digital Subscriber Line). Las otras dos posibilidades son Ethernet punto a punto sobre fibra óptica, que permite brindar servicios de Gigabit Ethernet (IEEE 802.3z) y 10 Gigabit Ethernet (IEEE 802.3ae), y Ethernet punto a multipunto sobre fibra óptica, lo que se conoce como EPON (Ethernet Optical Passive Network).
EoVDSL permite aprovechar la enorme cantidad de enlaces de cobre existentes en las redes de acceso, logrando velocidades de transmisión de datos de entre 5 Mbps y 15 Mbps. Esta tecnología, además, permite brindar servicios de voz, video y datos. El alcance de EoVDSL se encuentra, típicamente, entre 300 m y 1.5 Km.
Ethernet Punto a Punto permite brindar enlaces de fibra dedicados, dando lugar a lo que se conoce como FTTX (Fiber-To-The-X). Estos enlaces permiten que los usuarios tengan acceso a Internet a velocidades de transmisión muy elevadas, y con un ancho de banda y disponibilidad garantizados. Una ventaja interesante de la fibra óptica es su baja atenuación, que permite alcanzar distancias de más de 40 Km con este esquema.
EPON a diferencia de Ethernet Punto a Punto, permite brindar servicios de gran ancho de banda y elevadas velocidades de transmisión, pero a bajo costo. Con esta tecnología se pueden alcanzar velocidades de transmisión de más de 100 Mbps. EPON contribuye a bajar los costos de la red, dado que minimiza la cantidad de fibras a utilizar y la cantidad de equipos y transceivers necesarios. Con EPON se pueden alcanzar distancias de hasta 20 Km.
En la figura 1 se pueden observar las distintas topologías utilizadas para Ethernet en la primera milla –en el caso de utilizar fibra óptica–, y los dispositivos que deben ser utilizados, como los OCX (Optical Cross Connect) y las ONU (Optical Network Unit). La imagen superior de la figura muestra el modelo de Ethernet punto a punto sobre fibra óptica, para el caso de EoVDSL el enlace será de cobre en lugar de fibra y los equipos serán distintos. En la imagen inferior de la figura se puede apreciar un ejemplo de EPON.

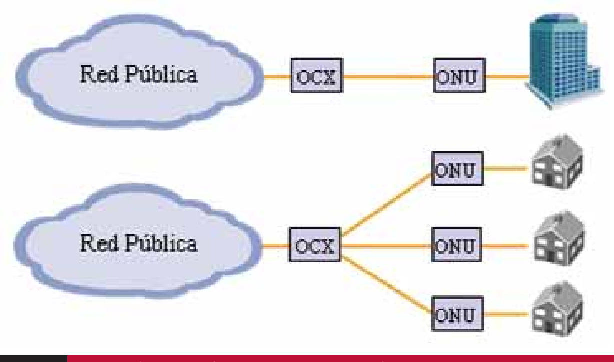


Fig.1 Topologías de Acceso por Fibra

Conclusión

Actualmente Ethernet soporta todos los servicios que un proveedor puede brindar –con una calidad de servicio adecuada–, como ser acceso a Internet de alta velocidad, Telefonía IP y Streaming de Video, con lo cual el uso exclusivo de otras tecnologías en las redes de acceso ya no es una necesidad.
Por otro lado, hoy en día el 99 por ciento del tráfico de datos que se transmite a través de una red WAN (Wide Area Network), comienza y termina como Ethernet e IP. EFM cierra la brecha tecnológica que permite tener una red enteramente basada en Ethernet, lo cual simplifica el diseño y mantenimiento de la red, y al estar basada en un mismo protocolo, se pueden diseñar y brindar soluciones globales sin la necesidad de realizar una conversión de protocolos. ●

Fundamentos de Networking

La Capa Física



La primera capa del modelo OSI se encarga de recibir los datos de las capas superiores y convertirlos en un flujo de bits, para luego enviarlos hacia el destino. Las funciones de esta capa están íntimamente relacionadas con los parámetros del medio físico.

Miguel F. Lattanzi
Ingeniero en Telecomunicaciones (IUPFA)
SIEMENS – Red de Acceso y Datos

Juan M. Urti
Ingeniero en Telecomunicaciones (IUPFA)

Nota #1 de 5 Serie

1- La Capa Física.

2- La Capa de Enlace: VLAN, Spanning Tree, Bridges y Switches.

3- La Capa de Red: Direccionamiento IP, Protocolos de Enrutamiento, Routers.

4- Tecnologías de WAN: Fundamentos de BGP, Concepto de VPN, VPN basadas en MPLS.

5- Management y IP QoS: SNMP, Prioridades y Reserva de Recursos, Manejo de Colas.

Introducción

Muchas veces la importancia de tener conocimientos sobre los medios de transmisión, los distintos tipos de conectores y los distintos códigos de línea –utilizados para codificar las señales de información–, es dejada de lado.
Uno tiende, a menudo, a perfeccionarse en los aspectos más avanzados de las capas superiores a ésta, como ser switching, routing o el management de las redes, pero muchas veces los problemas existentes, se encuentran relacionados con el medio de transmisión.
Entre los inconvenientes más frecuentes podemos mencionar, a modo de ejemplo, los siguientes: conectores defectuosos o mal conectados, cables mal armados, empalmes de fibra óptica mal hechos o equipos interconectados a mayor distancia que la soportada por el medio físico utilizado.
Los elementos definidos en la capa física del modelo de referencia OSI (Open System Interconnect), deben ser los primeros en ser verificados a la hora de presentarse inconvenientes de conectividad. Luego se deben revisar las funcionalidades de la capa de enlace de datos, de red, y así sucesivamente, hasta alcanzar las capas superiores del modelo.
Cabe aclarar, que los problemas de conectividad, pueden ser tanto por causas físicas como lógicas, generalmente estas últimas están ligadas a configuraciones erróneas. Es por ello, que se deben tener conocimientos relacionados con los medios de transmisión, para poder entender las posibles causas y encontrar la solución de la manera más eficaz.

Capas OSI

El modelo OSI esta basado en siete capas (inferior a superior): Física, Enlace de Datos, Red, Transporte, Sesión, Presentación y Aplicación. Es un modelo de referencia que ayuda a definir los medios de transmisión y sus parámetros, el alcance de los protocolos y servicios, y cómo estos deben interactuar entre los distintos niveles jerárquicos.

Medios de Transmisión

Por medio de transmisión se entiende al medio físico a través del cual es posible transmitir señales de datos, voz y video entre un extremo transmisor, el cual genera y codifica dichas señales, y otro extremo receptor, el cual las recibe y las decodifica para poder utilizar la información enviada.
Básicamente, los medios fisicos utilizados para la transmisión de señales de información, se clasifican en dos grandes grupos: medios guiados y medios no guiados.
Entre los medios guiados se encuentran el par de cobre trenzado UTP (Unshielded Twisted Pair) y STP (Shilded Twister Pair), el cable coaxial, la fibra óptica y las guías de onda. Entre los no guiados se encuentran el aire y el vacío.
Sabiendo esto, la diferencia entre ambos es evidente, en los primeros las señales viajan a través del medio y confinadas en el mismo; mientras que en el segundo caso, las señales se propagan a través del medio con total libertad, espacialmente hablando, debido al tamaño intrínseco de éste. Un ejemplo típico de los medios no guiados, es la transmisión de radio AM (Amplitud Modulada) o FM (Frecuencia Modulada).
Los medios de transmisión tienen, asociados a ellos, una infinidad de parámetros físicos, tipos de conectores y códigos, los cuales están normalizados por



medio de distintos organismos y son contemplados por el modelo OSI.
A continuación nombramos y definimos los estándares del IEEE (Institute of Electrical and Electronics Engineers) correspondientes a las distintas velocidades de transmisión para redes Ethernet (IEEE 802.3).

IEEE 802.3 –Ethernet– prevé una velocidad de transmisión de datos de 10Mbps. Los esquemas de cableado incluyen a 10BaseT (10Mbps en banda base). Esta pensado para entornos de LAN (Local Area Network).

IEEE 802.3u –Fast Ethernet– prevé una velocidad de transmisión de datos de 100Mbps. Entre los esquemas de cableado se incluyen 100BaseTX -dos pares de cobre-, 100BaseT4 -cuatro pares de cobre- y 100BaseFX -fibra óptica-. Este estándar define métodos de autonegociación para la selección de la velocidad de transmisión (10Mbps o 100Mbps). Esta pensado para entornos de LAN.

IEEE 802.3z –Gigabit Ethernet– prevé una velocidad de transmisión de datos de 1000Mbps (1Gbps). Los esquemas de cableado prevén 1000BaseSX (fibra óptica hasta 500m), 1000BaseLX (fibra óptica hasta 5km), 1000BaseT (par de cobre hasta 75m) y 1000BaseCX (par de cobre STP y cable coaxial hasta 25m). Este estándar es totalmente compatible con las instalaciones basadas en Ethernet y Fast Ethernet. Esta pensado para entornos de LAN y MAN (Metropolitan Area Network).

IEEE 802.3ae –10 Gigabit Ethernet– prevé una velocidad de transmisión de datos de 10Gbps para transmisión por fibra óptica. Los esquemas de cableado más comunes incluyen a 10GBaseSW (hasta 300m), 10GBaseLW (hasta 10km) y 10GBaseEW (hasta 40km). Esta pensado para entornos de LAN, MAN y WAN (Wide Area Network).

Parámetros

Las señales de información son afectadas por varios parámetros relacionados con los medios de transmisión. Estos son básicamente tres: ruido, atenuación y distorsión.
El ruido puede clasificarse en dos grandes grupos en función de cómo es generado. En el primer grupo -ruido interno-, el ruido se genera dentro de los equipos de comunicaciones, mientras que en el segundo -ruido externo- es una propiedad del medio de comunicaciones. Las causas más frecuentes de ruido interno se deben a los circuitos electrónicos de los equipos, tanto de sus elementos pasivos (cables y resistores) como de los activos (diodos, transistores y circuitos integrados). En cuanto al ruido externo, las causas son más diversas. Entre éstas podemos mencionar a la estática debido a tormentas eléctricas y al ruido cósmico -producido por la irradiación de las estrellas, como el sol- que afectan principalmente a las señales que se propagan por el aire y el espacio. Los pares de cobre, en cambio, son afectados mayormente por el ruido blanco generado por el movimiento aleatorio de los electrones en un conductor.
La atenuación es la perdida de amplitud de la señal, a medida que se propaga por el medio de transmisión. Para solucionar este problema se deben utilizar repetidores regenerativos, los cuales reconstruyen la señal y la retransmiten hacia el próximo repetidor o hacia el destino final. Los repetidores deben estar a una distancia tal que la amplitud de la señal esté por encima del umbral de regeneración. Si la amplitud es menor, muy posiblemente el repetidor no pueda regenerar la señal, lo cual produciría una perdida de información. La atenuación tiene mayor impacto sobre las señales analógicas, suele medirse en decibeles por metro (dB/m) o por kilómetro (dB/km). El tercer y último parámetro es la distorsión. Ésta produce deformaciones en la señal la cual degradan a medida que se propaga por el medio de transmisión. Si la distorsión es la suficiente como para deformar la señal por debajo del umbral de distorsión, ésta puede no ser interpretada en el destino o en los repetidores para poder ser regenerada. Además se corre el riesgo de una mala interpretación de la señal, lo cual puede producir cambios en la misma que alterarían la información que viaja asociada a ésta.
Los tres parámetros mencionados hasta ahora afectan a todos los medios de transmisión y tipos de señales -analógicas o digitales- en mayor o menor medida. Por otro lado tenemos una serie de parámetros, que afectan a las redes que se basan en la conmutación de paquetes -típicamente las redes de datos-, los cuales causan distintos efectos y tienen distintos grados de impacto en función de si se trata de comunicaciones de datos en tiempo real o no. Los mismos están definidos en el cuadro Parámetros que afectan la transferencia de Datos.

Relación Señal a Ruido

En los sistemas digitales el ruido aumenta la tasa de error BER (Bit Error Rate). No es importante el ruido por sí solo, sino el nivel de ruido en relación al nivel de la señal de información. Esta relación señal a ruido (S/N) se suele expresar en decibeles (dB) y mientras mayor sea su valor, mejor será la calidad de la señal de información.

Transmisión por Fibra Óptica

La fibra óptica es un excelente medio de transmisión guiado, es tan eficiente y seguro que hace tambalear, incluso, a las comodidades y los bajos costos de los medios no guiados. Tiene una atenuación mínima, casi no es afectada por la distorsión y es totalmente inmune al ruido. Además es muy segura, debido a que es muy difícil seccionar un cable de fibra para poder interceptar los datos transmitidos.

Topologías

La topología de una red de comunicaciones -tanto de datos como de telefonía-, define el esquema físico de interconexión que tendrán los distintos equipos que la forman. La topología esta ligada, fundamentalmente, con el método de acceso al medio y con las tecnologías de transmisión empleadas. En la figura 1 se pueden observar las topologías mas utilizadas.
La (a) utiliza un equipo en el cual se concentran todos los enlaces. Sus funciones principales son básicamente dos: procesar las señales entrantes para poder regenerarlas y repetirlas hacia su destino, y realizar las funciones de conmutación necesarias para que los distintos equipos se puedan comunicar entre sí. Esta topología es la más usada en redes LAN, el dispositivo concentrador puede ser tanto un Hub como un Switch.
En (b) se puede observar que todos los hosts están interconectados, este esquema además de ser más complicado que (a) es mucho más costoso. Los primeros sistemas telefónicos estaban interconectados de esta manera, hasta la aparición de las primeras oficinas de conmutación. En redes de datos es muy difícil encontrar un esquema como éste, pero en algunos casos donde es necesario garantizar un 100 por ciento de disponibilidad en los enlaces y se requiere una elevada redundancia física, se pueden observar cores –núcleos de red– con nodos interconectados de esta manera.
La topología (c) se utilizó mucho en los primeros años de las redes de datos. Todos los host –menos el de origen– recibían los datos enviados. En los extremos se debían colocar conectores terminales resistivos, para atenuar las señales que viajaban por el medio físico, logrando así que las mismas no rebotaran durante largos períodos de tiempo manteniendo el medio de transmisión ocupado.
La configuración (d) también fue muy utilizada en redes de datos LAN y aún hoy en día –en menor medida– se sigue utilizando principalmente en redes que no han migrado al estándar Ethernet. Las redes que utilizan Token Ring (IEEE 802.5) -desarrolladas por IBM– hacen uso de esta topología, en la cual cada uno de los host hace las veces de repetidor, capturando las señales de datos que le llegan, para procesarlas y reenviarlas hacia el siguiente host.
Esta topología, actualmente, es muy utilizada en entornos de MAN en redes ópticas SDH (Synchronus Digital Hierarchy) y SONET (Synchronus Optical Network) así como en las nuevas redes Metro Ethernet, que hoy en día se están instalando en todas partes del mundo.

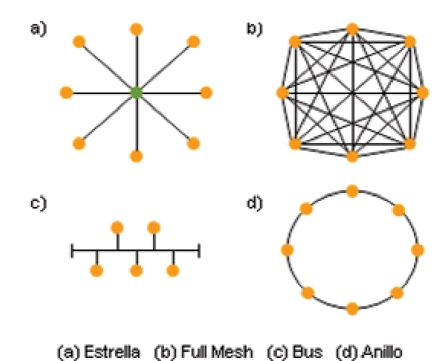


Fig.1 Topologías de Red

Parámetros que afectan la Transmisión de Datos

Delay
El retardo es producido, básicamente, por el tiempo extra que introducen los distintos equipos por los que pasan los paquetes, debido al procesamiento que realizan. Afecta principalmente al tráfico de datos para aplicaciones en tiempo real.

Jitter
La variación del retardo -o variación del delay- se produce cuando los intervalos de retardo introducidos entre los paquetes, no son constantes. Este efecto puede generar solapamiento entre paquetes -y la consecuente perdida de información- y problemas a la hora de reproducir información en tiempo real.

Packet Loss
La perdida de paquetes es producida principalmente por fallas de hardware, enlaces sobrecargados de tráfico y la degradación de la señal de datos en el medio de transmisión. Afecta en gran medida a las aplicaciones en tiempo real, dado que en una reproducción de audio o video no hay margen para la pérdida de datos, carece de sentido reproducir los mismos fuera de orden.

Tecnologías de Acceso al Medio

En el mundo de las redes de computadoras el acceso al medio, es la manera en que los host se turnarán para la utilización del canal de transmisión. Aunque hubo muchas propuestas por parte de los fabricantes, solo dos métodos se impusieron: Carrier Sense Multiple Access y Token Ring.
El acceso múltiple por detección de portadora –**CSMA**– es un tipo de mecanismo de características *probabilísticas* empleado en las redes Ethernet, para mantener un esquema de transmisión y recepción de tramas de forma ordenada y fiable, evitando al mínimo posible las colisiones y retransmisiones entre terminales.
La primera topología utilizada en las redes Ethernet era la de BUS compartido con control distribuido. Este canal de comunicación –el bus– era pasible de acceso por cualquier terminal para transmitir todo el tiempo, por lo que se precisó hallar la manera de que todas las computadoras envíen y reciban información a su debido tiempo, y sin superponerse. Para comprender el concepto de bus compartido debemos remontarnos a explicar brevemente cómo era la primera topología física, que dio origen a este mecanismo de acceso hoy ampliamente usado.
El diseño original se basaba en un cable coaxial totalmente pasivo, denominado "ether", al cual se conectaban todos los terminales de red. Las computadoras accedían a este por medio de un transceptor, el cual atravesaba ligeramente el blindaje del cable hasta encontrar el conductor central (otros fabricantes preferían cortar el ether e insertar un terminador en forma de "T"). El transceptor era capaz de transmitir y recibir señales eléctricas, provenientes o con destino el ether, y además poseía un cable denominado AUI (Attachment Unit Interface) que vinculaba al propio transceptor con la tarjeta de red del terminal final.

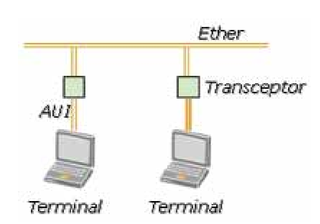


Fig.2 Modelo de Conexión Ethernet

Como es de imaginarse -y como hemos mencionado– este cable coaxial ha ido migrando paulatinamente a los hoy conocidos cables de par trenzado, y adicionalmente, el esquema de transceptores evolucionó a una solución en la cual un dispositivo central, denominado Hub, actúa como concentrador del cableado y de la red LAN.
A pesar de los cambios, la filosofía empleada para la transmisión no varió en su esencia, y desde aquel día hasta hoy la tecnología Ethernet emplea a CSMA/CD (Carrier Sense Multiple Access with Collision Detection) como método de acceso al medio.
Todas las computadoras conectadas al bus –o **ether**– , antes de comenzar una transmisión de una trama, deben censar el canal para ver si algún otro terminal lo está usando –en caso de que esté en uso se detectará una onda portadora–, y si no observa presencia eléctrica en el bus estará autorizado a transmitir información.
Como mencionamos anteriormente, el acceso es distribuido porque no hay una autoridad central encargada de otorgar los permisos para el uso del bus. A pesar de ello, los envíos de información están limitados en duración, ya

Confianza, Resultados, Clientes, Negocios, Flexibilidad ante el Cambio, Servicios Profesionales, Procesos de Negocios, Innovación.

Integrated Solutions !

Patagonia Technologies
Integrated Solutions

Produced by: www.e-synch.com.ar

Modern Design™
La respuesta inteligente para la próxima generación de soluciones SOA

http://mdesign.patagonia-tech.com

Alianzas



Microsoft











AMD

Nuestra propuesta de Valor



Business Solution Division
AMF - Pricing - Portal Banking - Orchestration XSource2SOA

Modern Design Division
"Next Generation Solutions: SOA Concept"
Sun

Mobile Solution Division

Contact Center Division

SAP Division

Nuestras "Killers" Aplicaciones

Portals

Framework de presentación con la capacidad de exponer interfaces de usuarios y asociarlas a servicios de Agentis. **Portals** ofrece interfaces prediseñadas en forma de elementos configurables. Todos los sitios y portales desarrollados con esta herramienta pueden integrarse a aplicaciones existentes y servicios Web.

AMF

Sarbanes Oxley (SOX) ha cambiado para siempre la forma de realizar una auditoria.
AMF es la aplicación más potente y segura del mercado que utilizando Agentes Inteligentes, Monitorea Logs bajo norma SOX Compliance.

XSource2SOA

Uno de los puntos más críticos en la vida de una compañía es la migración de sistemas. **XSource2SOA** es una aplicación que convierte lenguajes (RPG, Cobol, 4GL, Oracle Forms) a Java de manera simple, rápida y sencilla.



Tte. Gral J. D. Perón 949 Piso 12 / C1038AAS - Ciudad de Buenos Aires - Argentina
Tel: (5411) 4393-7227 / www.patagonia-tech.com / info@patagonia-tech.com

que la trama Ethernet 802.3-2002 (y sus versiones anteriores) posee un tamaño máximo en el campo información, evitando de esta manera que alguien tome el canal de forma indefinida (aquí nos referimos a Ethernet y 802.3 CSMA/CD como sinónimos, en el próximo número explicaremos una pequeña diferencia que las separa).

En este tipo de redes todos los host tienen acceso al medio todo el tiempo, pero solo se les permitirá transmitir de a uno a la vez ya que el medio es compartido.

En un medio de transmisión de cobre, las señales se propagan a una velocidad proporcional al 80 por ciento de la que se desplaza la luz. Por tal motivo un terminal puede detectar libre el canal e iniciar una transmisión, pero simultáneamente en el otro extremo de la red otro host puede realizar lo mismo, teniendo como resultado una **colisión**.

Las colisiones se producen cuando 2 o más señales eléctricas se superponen –se duplica la tensión de la señal–, perturbando la información que llevan con ellas. Aquí observamos porqué las primeras redes tenían un único dominio de colisión, debido a que cuando se desea enviar una trama a otro terminal, en realidad ésta se transmitía a todas las computadoras conectadas a la red –ya que todas deben saber si el bus está o no libre–, dejando propensa a todas las máquinas a producir una colisión si intentasen transmitir información alguna.

Cuando se detecta una colisión, los host esperan un tiempo aleatorio para reintentar transmitir. La estación debe interrumpir la transmisión del paquete y transmitir una breve señal de interferencia –denominada secuencia de jamming– para asegurar que todas las estaciones participantes en la colisión la detecten.

El tiempo además de aleatorio es exponencial –para ver más detalles acerca del algoritmo empleado, ver el apartado referencias– , ya que si detecta nuevamente otra colisión esperará el doble de tiempo, y así sucesivamente, reduciendo las probabilidades de hallar una falla al mínimo. Esta capacidad de detectar colisiones, es la que hace a las redes Ethernet propiamente CSMA/CD.

El peor caso de colisión se produce cuando los dos terminales que transmiten están físicamente lo más lejos posible -por norma-. Si la segunda estación comienza a emitir una trama infinitesimalmente antes de recibir el primer bit del emisor, debemos sumar los tiempos de propagación de la primera trama y de la segunda, al cual denominaremos "ventana de colisión".

Por lo tanto para asegurarnos de no tener colisiones indetectadas, las transmisiones deben durar más que la ventana -es por este motivo que la trama debe tener al menos 46 Bytes- y las estaciones deben chequear la ausencia de transmisiones por al menos ese tiempo.

CSMA/CA

Como alternativa a CSMA/CD, existe aunque muy poco empleada, CSMA/CA (Collision Avoidance), en la cual luego de transmitir una trama, el emisor espera un acuse de recibo del receptor para indicar que la transmisión fue exitosa. No fue tan difundida debido a que la red pierde eficiencia y ancho de banda real.

El otro método de acceso al medio es **Token Passing**, utilizado en las redes Token Ring. Es un método de acceso determinístico, dado que el sistema en sí, determina qué estación puede transmitir en un determinado instante de tiempo, en lugar de realizar un censado del medio. Vale aclarar que en las redes en anillo la información es transmitida en un solo sentido, siendo él mismo configurable. Este aspecto es muy importante para entender cómo funcionan las redes Token Ring y cualquier red –cuya topología sea en anillo– en general.

La explicación de funcionamiento que se desarrollará a continuación, puede verse gráficamente en la figura 2. En este ejemplo, el sentido de transmisión de la información en la red es anti-horario (A-B-C-D).

Mientras el Token circula por la red, ninguna estación esta transmitiendo o recibiendo datos. En el momento en que una estación decide transmitir información (A), espera a que arribe el Token, lo retiene y lo modifica para indicar la secuencia de comienzo de trama, le adjunta la información a enviar y retransmite toda la trama nuevamente hacia la red. A medida que la trama circula por la red, las estaciones intermedias (B) la revisarán, y al ver que no son el destino la reenviarán nuevamente. Cuando la trama llega a la estación de destino (C), ésta copia localmente la información, realiza una comprobación de errores y marca la trama para indicar que la misma llegó bien, luego de lo cual la retransmite. De ésta manera, la trama de información ya aceptada por el destino, continúa su recorrido hasta llegar nuevamente a la estación de origen –las estaciones intermedias (D) se comportarán como antes–.

Token
Trama de Datos (Origen : A, Destino : C)
Trama de Datos marcada con acuse de recibo (ACK)

Fig.3 Métodos de acceso Token Passing

Cuando la estación de origen recibe la trama verifica que la misma haya llegado bien al destino, si ese es el caso quita la información de la trama y la descarta. Luego modifica el Token para dejarlo con sus valores originales y lo transmite hacia la red, para que comience a circular nuevamente. En caso de que un host quiera realizar una transmisión de datos, todos los pasos anteriores se repetirán.

HUBs

Un Hub en un dispositivo de red de capa 1 que actúa como concentrador del cableado de una red LAN, uniendo en un bus físico interno a todos los terminales de red. Todas las computadoras comparten el mismo medio y, por consiguiente, se encuentran en el mismo dominio de colisión, teniendo una permanente contienda por el acceso al medio y el ancho de banda. La única diferencia entre las primeras LAN y estas, es que las más modernas poseen una topología en estrella y ya no bus.

Un Hub no trata el tráfico que atraviesa la red y, por ende, tampoco la totalidad de sus puertos, ya que solo se encarga de extender el medio físico y repetir la señal recibida por un puerto por todos los demás. Muchos hubs adquirieron funcionalidades de capa 2, pero en este artículo solo haremos referencia a los puramente de capa 1, que solo se encargar de repetir y regenerar la señal.

Brevemente debemos mencionar, como observamos en la figura 4, que cuando un terminal desea transmitir información a otro host, todos los demás terminales de la red reciben la trama –aquellos que no sean destino eliminarán la trama– debido a que hay un único dominio de colisión. ●

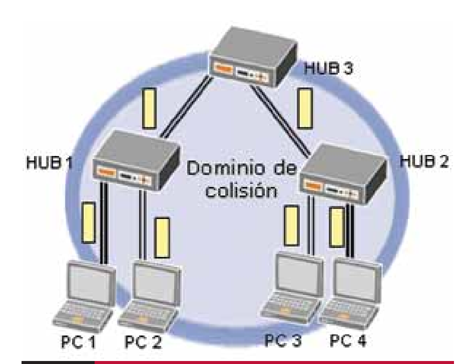


Fig.4 Esquema de red con HUBs

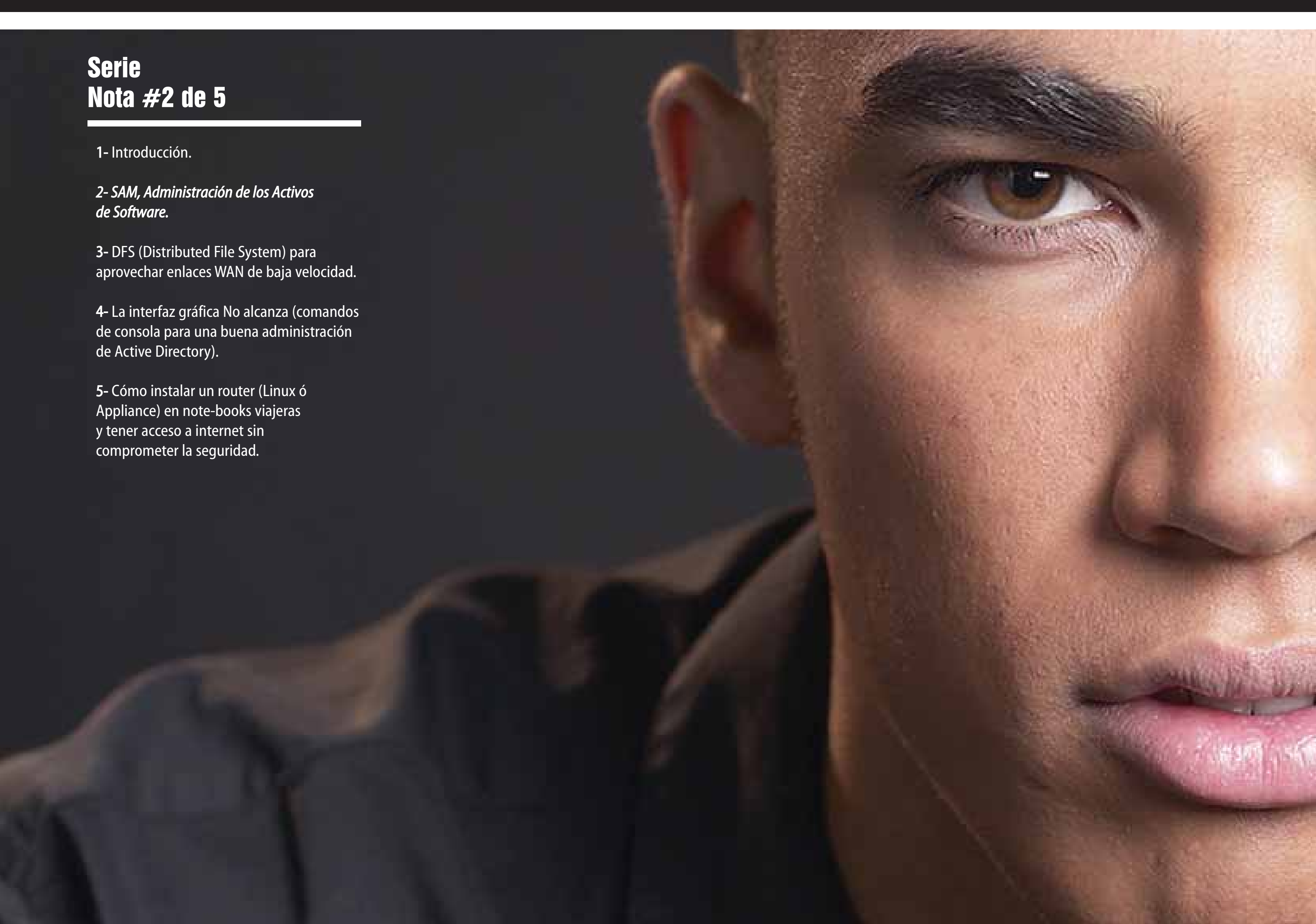
Lectura Adicional

[1] Tanenbaum A. Redes de Computadoras. Pearson, 4ta. Edición, 1997.
[2] Blake R. Sistemas Electrónicos de Comunicaciones. Thomson, 2da. Edición, 2004.
[3] http://www.cisco.com
[4] http://www.ieee.org

IT Pro en PyMEs

Por **Sebastián Passarini**
Administrador de Redes

SAM es, además de una herramienta, una guía de buenas prácticas, la cual permite al departamento de IT lograr todos los objetivos en cuanto a la administración de activos de software se refiere.



Serie
Nota #2 de 5

1- Introducción.

***2-** SAM, Administración de los Activos de Software.*

3- DFS (Distributed File System) para aprovechar enlaces WAN de baja velocidad.

4- La interfaz gráfica No alcanza (comandos de consola para una buena administración de Active Directory).

5- Cómo instalar un router (Linux ó Appliance) en note-books viajeras y tener acceso a internet sin comprometer la seguridad.

Administración de los Activos de Software - SAM

Uno de los desafíos que enfrenta el departamento de IT es lograr mantener el orden y el control de los activos de software de la compañía. Como IT Pros de tecnología informática, será necesario contar con los conocimientos y herramientas que nos ayuden a cumplir esta tarea. Si bien el objetivo principal es saber si estamos o no en regla con nuestras licencias, veremos a lo largo de este artículo que una buena administración de la información colectada (nuestro inventario) nos ayudará, entre otras cosas, a descubrir un software obsoleto el cual, una vez reemplazado, agregará mayor valor y mejorará los resultados de los procesos del negocio. En el recuadro Riesgos y Beneficios, podemos apreciar las consecuencias que se desprenden de la realización o no de un plan SAM.

El proceso de administración de activos de software comprende un ciclo permanente, del que podemos destacar las siguientes etapas:

- **Inventario de software.**
- **Confrontar inventario y licencias.**
- **Políticas y procedimientos.**

Inventario de Software: debemos examinar cada una de las PCs de la compañía en búsqueda del software instalado. Esto podemos hacerlo, como veremos más adelante, en forma manual, automática o mediante la combinación de ambas, dependiendo nuestra elección de los recursos con los que contamos y del tamaño/disposición física de nuestra red. Al finalizar nuestro inventario inicial será muy importante poder consolidar la información en, por ejemplo, una planilla Excel.

Confrontar inventario y licencias: antes de poder realizar esta tarea es imprescindible contar con un archivo físico donde podamos almacenar en forma de hard-copy (copia impresa) toda la documentación de respaldo correspondiente al software de la compañía. Luego podremos confrontar nuestro inventario con las licencias disponibles, obteniendo así un resultado o estado de licencias. Finalmente, y en base a estos resultados, podremos tomar decisiones tales como la actualización del software obsoleto, la exclusión de determinado software de las PCs porque no se utiliza, permitiendo así la redistribución del mismo y evitando la adquisición innecesaria de nuevas licencias.

Políticas y procedimientos: debemos contar con políticas y procedimientos de adquisición de software que nos indiquen, por ejemplo, cómo y dónde almacenar la documentación de respaldo, quién es el encargado de la instalación de los mismos, quién es el encargado de aprobar la compra de un nuevo activo de software, cómo se da de baja un software obsoleto. Dichos procedimientos deben incluir un glosario de términos utilizados, como por ejemplo del término DSL ó Definitive Software Library, definición que ITIL (Information Technology Infrastructure Library) le da al repositorio de todas las copias maestras del software.

Beneficios

- **Elimina el riesgo de sufrir acciones legales por utilizar software sin licencia y evita que la compañía deba enfrentar costos inesperados.**
- **Ayuda a los directivos a comprender el verdadero costo del software al tener que renovar y adquirir nuevas licencias.**
- **Ayuda a controlar gastos mediante la redistribución de software instalado y no utilizado (uso de software).**

Riesgos

- **Multas monetarias por la utilización de software sin licencia que lo respalde, lo que también atenta contra la reputación y la imagen de la compañía.**
- **Problemas de incompatibilidad de software: el software obsoleto no es compatible con documentos realizados en versiones modernas del mismo, por otro lado tampoco contamos con soporte por parte del fabricante para la solución de problemas. Esto perjudica la comunicación y el desempeño de los usuarios, así como también la de los administradores.**

A continuación, y ahora sí como eje de nuestro artículo, veremos algunas de las herramientas y métodos disponibles para llevar a cabo nuestro inventario de software.
Conforme vayamos avanzando en la demostración de las herramientas (de lo manual a lo automático), veremos cómo tiende a aumentar el costo financiero y cómo a disminuir el operativo. El método a elegir dependerá más que nada del presupuesto con el que contemos. Lo ideal: gastar lo menos posible, mantener lo menos posible.

Inventario Manual: una forma clásica de colectar información acerca del software instalado en las PCs de la compañía es recorrer equipo por equipo y asentar en una planilla de Excel todo el software que encontramos instalado. El problema se presenta cuando tenemos decenas de equipos ya que el tiempo invertido en visitar cada puesto de trabajo, colectar la información y finalmente volcarla a una planilla de consolidación de datos, puede ser una experiencia muy frustrante como auditores de software. Además corremos el riesgo de no asentar correctamente todo el software instalado. Por suerte contamos con los comandos ***nbtstat.exe y reg.exe*** del sistema operativo de red, los cuales nos permiten mediante su ejecución dentro de un archivo batch, colectar esta información desde la consola del administrador. Para tener un detalle de cómo funciona esto, pueden acceder a http://windowsitpro.com/Article/ArticleID/41505/41505.html, o bien pueden descargar el archivo directamente desde http://emea.windowsitpro.com/Files/07/41505/Listing_01.txt
Solo un comentario: para que el archivo batch funcione, deben ordenar las líneas de código tal cual se muestran en la **Figura 1.**
Pero aún así tengamos todo el software de las PCs en un archivo de texto, el cual podemos abrir y procesar con Excel, nos queda contabilizar las licencias y seleccionar solo aquellas aplicaciones de las que nos interesa conocer su estado. Con esta solución estamos colectando desde aplicaciones hasta service packs y parches del sistema operativo. Solo imagínense lo que eso significa. Mucha información irrelevante. Ahora bien, hasta aquí solo hemos ahorrado dinero y la ardua tarea de movernos hasta las PCs a auditar, lo que no es poca cosa, pero ¿cuál es el costo operativo? Sin duda es muy alto con lo que no es recomendable como solución definitiva para nuestro plan SAM. Aunque podemos utilizarla como herramienta para brindar soporte remoto a los usuarios de nuestra LAN, ya que rápidamente tendremos en pantalla todo el software instalado en una PC determinada. Esto nos abre paso a una herramienta más confiable la cual además de ser gratuita disminuye considerablemente el costo operativo de la recolección de software.

```
If %ERRORLEVEL% EQU 0 Echo %targetcomp% is unavailable at this time & Goto :EOF

For /f "tokens=1,2,*" %%i in ('Reg Query
\\%targetcomp%\HKLM\Software\Microsoft\Windows\CurrentVersion\Uninstall
/S ^| Find /I "DisplayName" ^| Find /V /I "Quiet"') Do Call :getappname "%%k"
```

Fig. 1

MSIA – Microsoft Software Inventory Analyzer: en su esfuerzo por brindar soluciones a sus clientes y ayudar a la implementación de un plan SAM, Microsoft ha puesto a disposición esta herramienta, la que podemos descargar en forma gratuita y en varios idiomas desde su sitio web. Es de fácil operación y nos brinda un inventario con todo el software de Microsoft instalado en nuestras PCs, desde aplicaciones hasta sistemas operativos y sus service packs si así lo deseamos. Las opciones de búsqueda son variadas ya que permite realizar un inventario de software local o a través de la red y, dentro de esta última, cuenta con una selección de equipos desde el árbol de Dominio o ingresando un rango de direcciones IP. Facilita la consolidación de informes realizados en un mismo equipo o en otros equipos, permitiendo de esta manera la unificación de datos aún de sitios remotos. También cuenta con una gran lista de productos disponibles para auditar, en la que se encuentra la mayoría de los títulos del fabricante y, una de las características más interesantes, es que permite la confección de informes en formato HTML, Excel o en archivo de texto, los que pueden ser enviados vía e-mail por la aplicación.
Podríamos decir que es la herramienta ideal para un control básico de licencias Microsoft y, al ser gratuita, no podemos exigirle funcionalidades tales como la de informar el uso de software, o llevar un control sobre la caducidad de los contratos de licencias. Es un recurso muy valioso al alcance de todos y, solo resta mencionar luego de nuestro lab, algunos puntos a tener en cuenta para quienes decidan utilizar este producto. Las pruebas fueron realizadas sobre dos equipos con la misma configuración de hardware y software de base, uno con la versión en español y otro con la versión en inglés del MSIA, derivando en los siguientes resultados:

• El producto es instalado en "C:\Archivos de programa\ Microsoft Software Inventory Analyzer". Dentro de esta ruta tenemos la carpeta "Informes" o "Reports" según el idioma instalado, donde se almacenan los informes generados por la aplicación. También encontramos la carpeta ForWin9x, que contiene scripts para escanear equipos con sistemas operativos Win9x y Me.
• El 90 por ciento de los intentos por escanear todas las PCs de la red al mismo tiempo (100 en total) fueron interrumpidos por fallos inesperados del software, provocando el cierre de la aplicación. Como solución alternativa se pueden correr escaneos en grupos de 10 PCs, lo cual elimina el problema y al ser un proceso bastante rápido no genera mayores inconvenientes.
• Notamos que la instalación en español detectaba un software instalado, lo que la versión en inglés no. Al cotejar luego las bases (archivos "Products.txt" y "Reg.txt") comprobamos que la versión en español tenía más información que la versión en inglés.
• El producto solo reconoce software de Microsoft. Aunque si modificamos los archivos "Products.txt" y "Reg.txt" podemos lograr que la aplicación detecte software de terceros.

Operating Systems by Platform
Reports and options
WinXP
Unknown
WinNT
Win2000
Win2003

Fig. 2

Centennial-Discovery: entrando en una gama más alta de productos tenemos esta aplicación la cual permite auditar en forma completa todos los equipos de la red, colectando tanto el software como el hardware instalado. Reconoce todos los dispositivos que posean una dirección IP como switches, access points, impresoras de red, routers y mucho más, brindando información valiosa sobre cada uno de ellos. Es un producto que se licencia por cliente y cuenta con dos interfaces de acceso, una es la GUI o interfaz gráfica de usuario (interfaz de configuración – solo para el administrador) y la otra es una poderosísima interfaz web llamada Discovery Dashboard, desde donde podemos consultar todo tipo de reportes (ver figura 2). Trabaja con una base de datos MSDE o SQL y una vez implementada la solución no debemos hacer mucho más, ya que los agentes instalados sobre equipos Windows o Linux se encargan de enviar al servidor toda la información de auditoría. Es realmente una herramienta muy completa que, además de brindar información sobre el saldo positivo o negativo de las licencias (ver figura 3), nos indica si una aplicación es utilizada o no en un equipo y con qué frecuencia, lo que nos permitirá desinstalar software de un equipo donde no se utiliza para poder instalarlo en otro que sí lo necesita, evitando de esta manera la compra de licencias adicionales. Cuenta con una infinidad de recursos que nos ayudarán a completar con éxito la implementación de nuestro plan SAM y, dado que es imposible mencionar en este artículo todas sus virtudes los invito a visitar el sitio http://www.centennial-software.com, del cual podrán obtener toda la información necesaria y una versión Trial para realizar sus propios labs. Luego desde www.centennialonline.co.uk, podrán acceder al producto on-line.

Conclusión: auditar nuestros activos de software puede tornarse en una tarea simple o compleja, dependiendo por lo general de los conocimientos y recursos disponibles. Siempre debemos partir de un orden inicial, una planificación que ayude a comprender el objetivo buscado, conociendo ante todo los riesgos y beneficios existentes. Ese es el espíritu de un plan SAM, el cual debemos mantener mas allá de la solución que decidamos implementar.

Para mayor información: esto ha sido solo una introducción al tema inventario de software y nuestra capacitación no estará completa hasta que hayan ingresado al siguiente sitio, el cual posee información sumamente valiosa respecto al tema -No dejen de consultarlo! http://www.microsoft.com/latam/softlegal/sam/default.mspx ●

Manufacturer	Product Name	Total Copies	Used	Unused	Licenses	License Required
Microsoft Corporation	PowerPoint 95	2	0	2	1	1
Microsoft Corporation	PowerPoint 97	4	0	4	3	1
Microsoft Corporation	PowerPoint Office Professional Edition 2003	1	0	1	2	-1
Microsoft Corporation	PowerPoint Office XP Professional	6	1	5	6	0
Symantec Corporation	Norton AntiVirus 10	6	0	6	5	1
Symantec Corporation	Norton AntiVirus 2004	3	3	0	4	-1
Symantec Corporation	Norton CleanSweep Fast & Safe Cleanup 7	3	0	3	3	0

Fig. 3

Acerca de Sebastián Passarini

Sebastián Passarini es Administrador de Redes y Tecnología en una empresa del rubro plástico. Durante sus 10 años de experiencia en el área informática, ha tenido la oportunidad de trabajar en Hewlett Packard Arg, entre otras empresas. Ha realizado varios cursos de tecnología y ha obtenido la certificación MCP de Microsoft en Windows NT Server. Es actualmente estudiante de la carrera de Sistemas y espera graduarse el próximo año.

Suscribite
y ahorrá un 40%

Pagá menos y no te pierdas ninguna Edición…

Promo
Suscripción Anual
12 Ejemplares
$70

Promo
Año y Medio
18 Ejemplares
$95

¿PORQUE SUSCRIBIRTE POR AÑO Y MEDIO?

- 18 Ejemplares NEX IT en tu domicilio, sin costo de envío.
- Newsletter Mensual con las últimas novedades del mundo IT.
- Acceso a los contenidos exclusivos de nuestra Web .
- Web Hosting Dattatec FREE por un año, 100MB de espacio, 8 GB de trans.
- Antivirus Panda Internet Security Platinum 2006 FREE por 6 meses.

Para aprovechar estas promociónes contactate a:

suscripciones@nexweb.com.ar | www.nexweb.com.ar | +54 (11) 5031.2287/88

SUSCRIBITE AHORA Y ACCEDÉ A LOS CONTENIDOS DE NUESTRA WEB







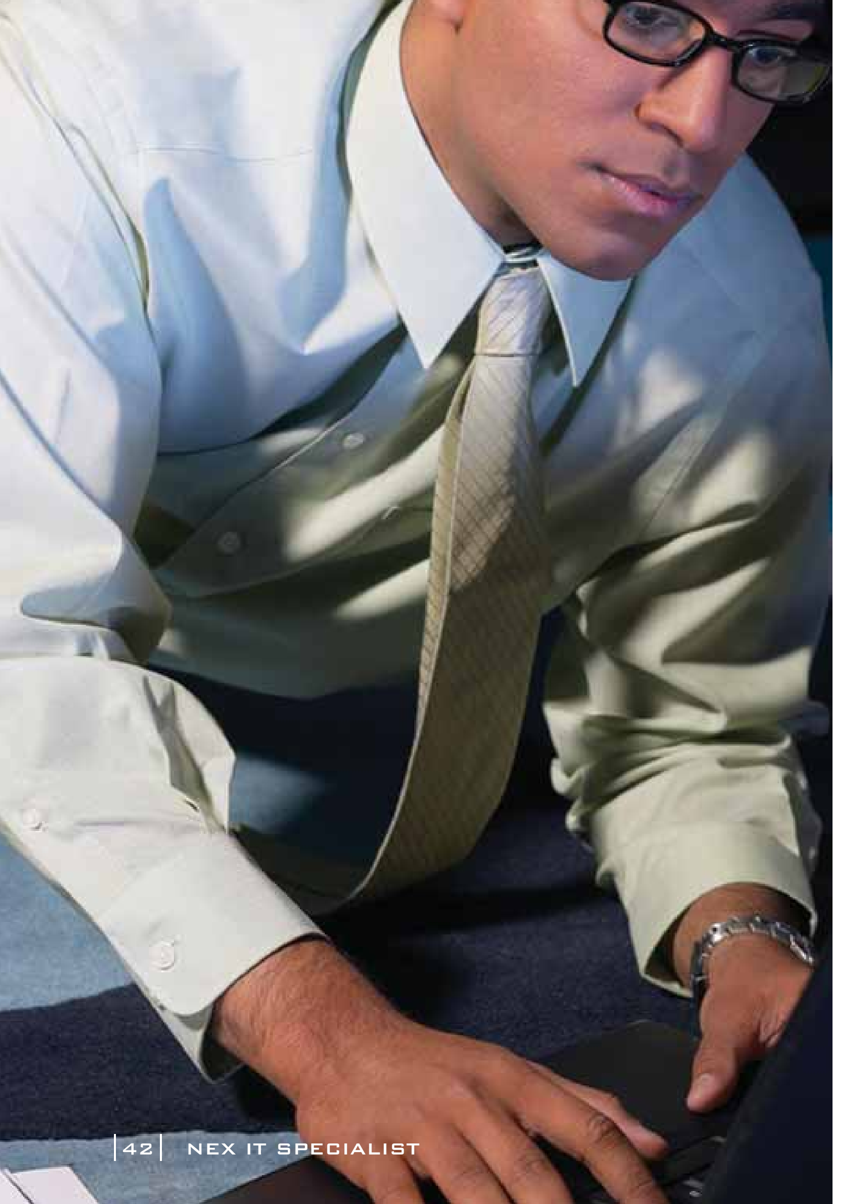


Terminal Server

Descubriendo lo nuevo.

Por: **Gustavo Gurmandi** MVP–MCT–MCSE-CCEA

En este artículo descubriremos las nuevas características de Terminal Server sobre la versión Beta2 de Windows Server Code Name "Longhorn". Veremos una introducción a los nuevos componentes que permiten hacer de Terminal Server una solución de acceso a aplicaciones propia.



Desde Windows NT 4.0 Terminal Server hasta Windows Server 2003, Microsoft ha ido ampliando las funcionalidades y mejorando la performance de los Servicios de Terminal Server. Pero con la nueva versión de Windows Server Code Name "Longhorn", Microsoft logra un cambio importante.

Hoy en día ya con la versión Beta 2 de Windows Server Code Name "Longhorn" podemos ver grandes novedades, donde los servicios de Terminal Server dejan de ser una funcionalidad para extender con herramientas de terceros y pasan a ser una solución de despliegue de aplicaciones propia.

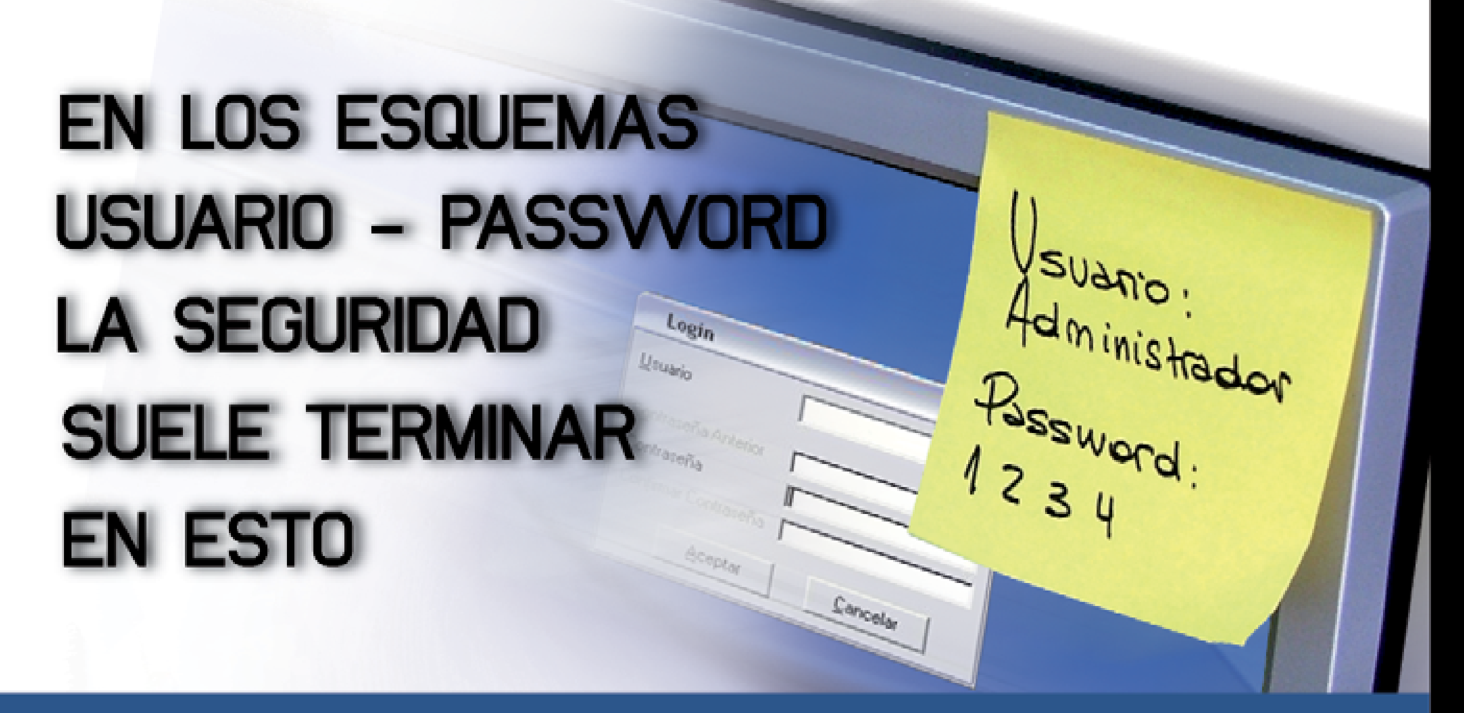
A través de sus nuevos componentes el Terminal Server de Windows Server Code Name "Longhorn" ofrece la novedad de acceder a **aplicaciones remotas** (uno de los grandes cambios) donde el usuario se conecta a una aplicación que esta instalada en el Terminal Server y solo visualiza la parte de la aplicación del Terminal Server (sin un desktop por detrás). Esto logra integrar la aplicación remota con el escritorio del usuario. Lo bueno de esto es que el usuario a simple vista no nota que está trabajando de forma remota vía Terminal Server.

En la **Figura 1** vemos un ejemplo de la integración de MSPaint Remoto en el Desktop de un usuario con Windows Vista.

El acceso a las aplicaciones remotas de la granja de Terminal Server se puede hacer desde el componente de **TS Web Access** o a través de la creación e instalación de archivos .rap.msi (100k aprox.) los cuales pueden ser manejados por software de Administración como SMS Server o a través de Políticas de Grupo de Active Directory para instalarlos a los clientes.



Fig.1



Autenticación *FUERTE* mediante dos factores "*algo que tengo*" (la llave) + "*algo que sé*" (un número de PIN)

Integre el manejo de los token USB a su sistema para aumentar la seguridad en el proceso de autenticación de usuarios.

El proceso es muy simple y se proveen ejemplos en distintos lenguajes de programación.

Autenticación de Administradores y Usuarios

En este caso se entrega una llave electrónica HARDkey MIO a cada usuario.
Esta llave es de uso personal y lo habilita a ingresar en la aplicación. Cada llave tiene asociado un número de PIN que el usuario puede configurar para evitar que un tercero pueda usarla en el supuesto caso que la pierda o se la roben.

Validación de accesos a Páginas Web

Este modo de protección permite que una página web publicada en un servidor de Internet, valide la identidad de los clientes que se conectan al mismo, utilizando como "clave de acceso" la llave electrónica más el PIN de la llave que debe conocer el usuario.



Bartolomé Mitre 777 2º A
(C1036AAM) Buenos Aires, Capital
Tel./Fax (54 + 11) 5500-7770/73

www.sitepro-sa.com.ar - info@sitepro-sa.com.ar

TS Web Access es el componente de Terminal Server que crea un sitio web sobre Microsoft Internet Information Services (IIS) 7.0, por el cual el usuario podrá ver sus aplicaciones disponibles en la granja de Terminal Server vía Internet Explorer y, con un solo clic, acceder a ellas de forma remota, tanto en la red local como desde Internet.
Si accedemos a las aplicaciones desde Internet deberíamos usar TS Gateway Server y así lograr una seguridad SSL (TLS 1.0) para la conexión http y para RDP.
Además existe la posibilidad de usar una Webpart de TS Web Access para integrar las aplicaciones de la Granja de Terminal Server con Windows Sharepoint Services o Sharepoint Portal Server. Los clientes soportados de TS Web Acces (beta 2) son: Microsoft Windows XP Service Pack 2, Microsoft Windows Server 2003 Service Pack 1, Windows Vista y Windows Server "Longhorn".
En la **Figura 2** vemos un ejemplo de TS Web Access y las Aplicaciones Remotas disponibles.
El otro método de acceso a las aplicaciones de la granja es por la creación los archivos .rap.msi de cada aplicación remota que queramos de la granja. En la creación de los archivos de aplicaciones remotas podemos elegir opciones de crear accesos directos a la aplicación remota en el escritorio del usuario o en menú de programas del cliente. También podremos integrar las extensiones de archivos de la máquina cliente con las aplicaciones remotas (por ejemplo enlazar las extensiones .docx con el Word 2007 instalado en la granja de servidores Terminal Server).
Otro avance importante con respecto a la seguridad en el acceso a las aplicaciones remotas o a los Terminal Server es el uso del componente de TS Gateway Server. Este se encarga de la seguridad de la conexión con la granja de Servidor de Terminal Server y su función principal es armar una conexión SSL (TLS 1.0) entre el cliente y él. Dentro del protocolo SSL (TLS 1.0) se encapsulara el protocolo RDP (HTTP Transport Layer Security (TLS) tunnel), lo que permite que la conexión al Terminal Server sea encriptado y así eliminar la necesidad de usar VPN para estos casos.
También desde la consola de TS Gateway podremos armar las reglas de acceso a la granja de Terminal Server basadas en **Connection Authorization Policies (CAP)** que pueden estar en un Central Network Policy Server o local en el Server de TS. Estas reglas nos permiten manejar qué grupos de usuarios o máquinas se podrán conectar a TS Gateway y qué tipo de dispositivo tendrán habilitado en el protocolo RDP (disk drivers, Printers, Serial Port, Clipboard...). También deberemos crear reglas de **Resource Authorization Policies (RAP)** que nos permitirán manejar qué recursos internos (servidores de TS) estarán disponibles para cada grupo de usuarios.
Otra de las funcionalidades de TS Gateway es que no solo nos podrá conectar con la Granja de Terminal Server sino con cualquier Sistema Operativo que soporte RDP (Windows Server 2003, Windows 2000 Server o Windows XP).
Otra de las mejoras en esta nueva plataforma es que estará soportado el uso de Aero Glass para las aplicaciones remotas que son ejecutadas desde los clientes. Para que esto se cumpla, ambos cliente y server deberán tener activado Aero Glass y usar RDP 6.0 con 32 bit de colores.
Se podrá dar prioridad a algún tipo de trafico dentro del protocolo RDP, donde podremos diferenciar los canales de tráfico de visualización y otros canales como clipboard, transferencia de archivos, o trabajos de impresión, donde a través de un cambio en el registro (en la versión Beta2) podemos poner una relación 70:30 dando más prioridad a la visualización que a los trabajos de impresión ya que estaría tomando solo el 30 por ciento para la impresión y demás canales. Obviamente también podremos adaptarlo a la necesidad de cada caso. Con el uso del último cliente de RDP, podremos obtener una resolución total de 4096 x 2048 teniendo en cuenta que ahora el cliente RDP soporta múltiples monitores en el cliente, utilizando la opción de mstsc.exe /span. En la Figura 3 vemos un ejemplo de Conexión con el cliente Remote Desktop 6.0
Los usuarios de Windows Vista podrán usar Single Sign-On con las aplicaciones remotas de la Granja de Terminal Server. La única configuración adicional que se debe hacer es la opción de **Allow Delegating Default Credentials** vía GPO de Active Directory o modificando la política local.
En cuanto a la compatibilidad de aplicaciones, podemos mencionar que se está optimizando la plataforma para que las aplicaciones Windows Presentation Foundations (WPF) puedan ejecutarse de manera transparente y con un rendimiento óptimo en un ambiente de Terminal Server.
Pensando que es la versión Beta 2 y que podemos esperar aún algunas mejoras en sus componentes dentro de la arquitectura de Terminal Server y la compra de Softricity, Microsoft gana terreno en el mundo de SBC (Server-Based Computing). ●

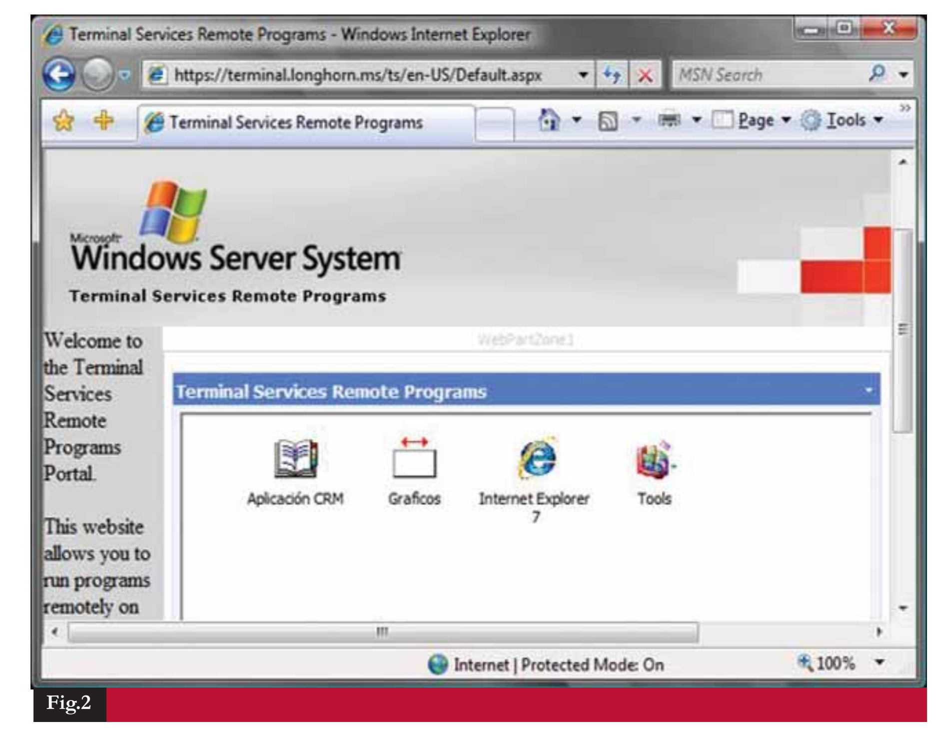


Fig.2

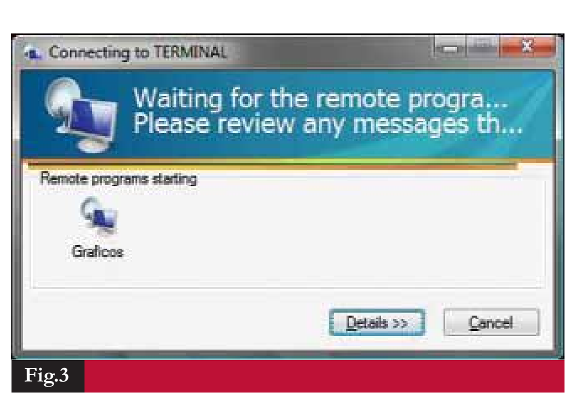


Fig.3

Sobre el Autor

Gustavo Gurmandi es especialista en Tecnologías Microsoft desde hace 8 años, es Certificado en Microsoft (MCSE y MCT), en Citrix (CCEA), ha obtenido el reconocimiento de Microsoft MVP en Terminal Server. Actualmente trabaja como Consultor y es activo colaborador de las actividades de Microsoft TechNet en Argentina.

ITIL: ¿En qué fase estamos?



Ricardo D. Goldberger

Periodista Científico especializado en Informática y Nuevas Tecnologías.

La historia de la informática está llena de ejemplos en los que metodologías, tecnologías y modelos pasaron de ser la panacea universal a grandes fracasos y desperdicios de inversión, casi sin solución de continuidad. En nuestra columna de hoy nos preguntamos en qué etapa de esa evolución se encuentra ITIL.

En la industria IT las siglas, y los significados a los que aluden, se transforman rápidamente en modas que, pasado un tiempo variable, se vuelven obsoletas para ser reemplazadas por otras. Algunas de esas siglas, especialmente cuando se refieren a metodologías o técnicas para enfrentar ciertas situaciones, suelen atravesar por tres fases:

1. La fase de la panacea. Ese objeto / modelo / tecnología / metodología sirve para todos los casos y para todo uso. Se organizan foros y conferencias para explicar la importancia de esa panacea, la razón por la cual todos debemos empezar a seguir sus indicaciones y, sobre todo, los argumentos para que invirtamos una buena cantidad de nuestro dinero (o el de nuestros inversores) en esta nueva panacea.

2. La fase de la desilusión. Cuando se descubre que hay una buena cantidad de situaciones en las que esa panacea se revela que no es tal, se cae en el extremo opuesto. En realidad no sirve para nada, el dinero que invertimos se desperdició totalmente, las herramientas que compramos quedan arrumbadas en alguna oficina deshabitada o en el fondo de algún cajón.

3. La fase del realismo encontrado. Repentinamente alguien recuerda que esa implementación que se hizo en el momento de la burbuja todavía está en funcionamiento... Y tan mal no nos fue. O después de una discusión, debate o investigación acerca de la mejor manera de encarar un proyecto se descubre que aquella tecnología resultó ser la mejor alternativa para ese caso en particular. Finalmente, ese objeto / modelo / tecnología / metodología encuentra su lugar, que no es el de la panacea universal, pero sí el de ciertas situaciones para las cuales resulta ser la mejor elección.

El mejor ejemplo de esto es el CRM. A fines de la década del '90 hubo un brote de congresos, foros y visitas de gurúes que nos explicaban que todas las empresas tenían que cambiar sus metodologías de trabajo, convertir al cliente en el foco y, sobre todo, invertir en toda esa cantidad de tecnología de nombres extraños que recién empezábamos a deletrear como data warehousing, contact center y, sobre todo, customer relationship management. Information Technology Infrastructure Library, cuya sigla ITIL se está haciendo cada vez más conocida, comenzó a recorrer ese camino.

En síntesis ITIL se define como una colección de libros que contienen guías para diferentes aspectos de mejores prácticas de la administración de los data centers.

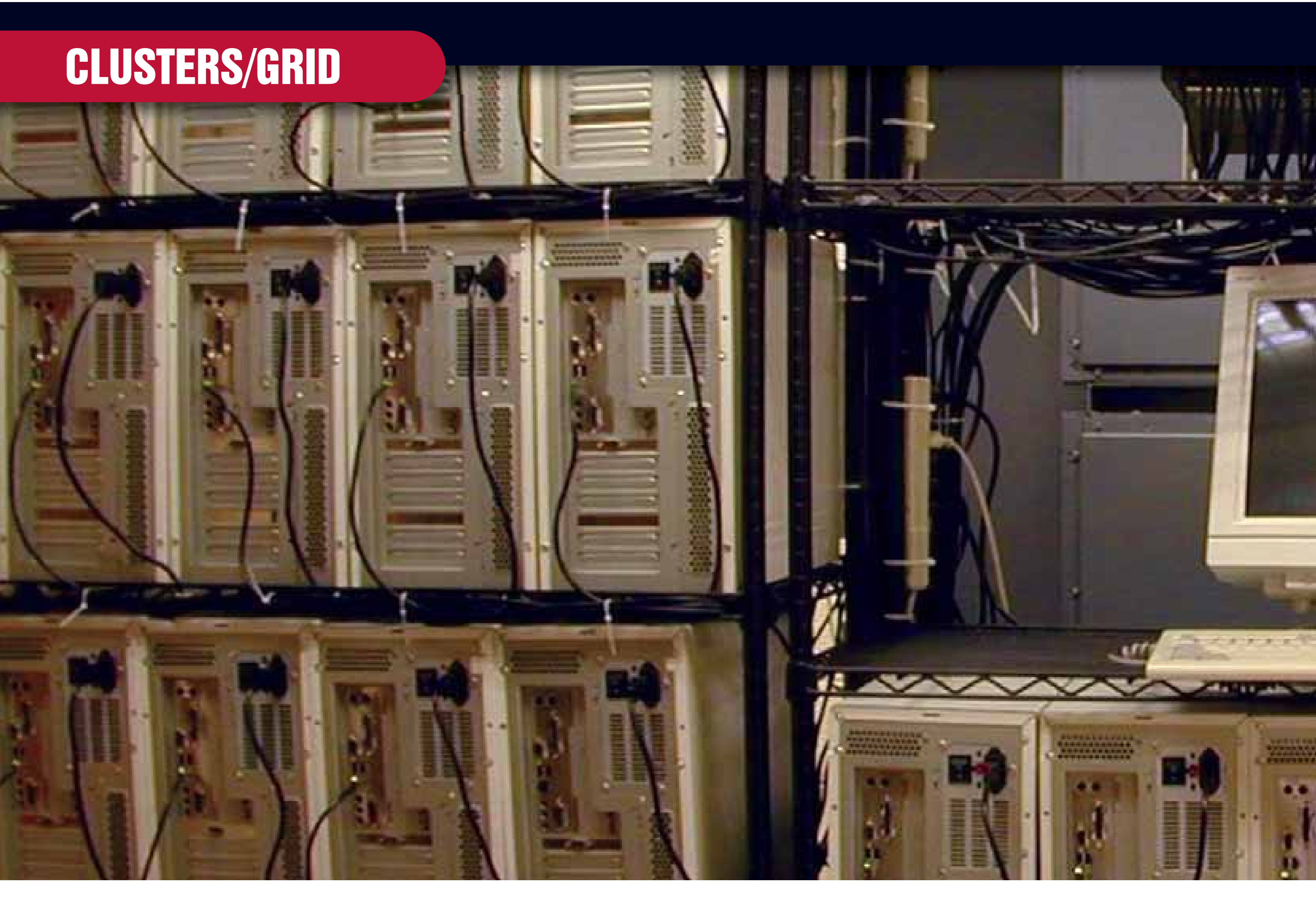
A pesar de que sus primeros escarceos pueden remontarse a mediados de la década del ´80, no es sino en mitad de los '90 que ITIL se convierte en un estándar para la administración de servicios IT del mismo tipo que la Information Services Procurement Library o la Application Services Library.

Todavía no estamos seguros de en qué fase de aceptación se encuentra ITIL, pero lo cierto es que oímos hablar y leer cada vez más acerca de estas metodologías.

Mi consejo es que se informe, trate de recabar la mayor cantidad de datos posibles como para que usted mismo pueda tomar una decisión. No creo que sea muy difícil pasar raudamente a una tercera fase, sin siquiera rozar la primera y la segunda. Quizá sea lo deseable. ●

Lectura Adicional

- **Itil Community Forum** (http://www.itilcommunity.com/index.php)
- **The IT Infrastructure Library** (http://itil.technorealism.org/)
- **The IT Service Management Zone** (http://www.itil.org.uk)



High Performance Computing a un costo reducido

Clusters Beowulf bajo Linux

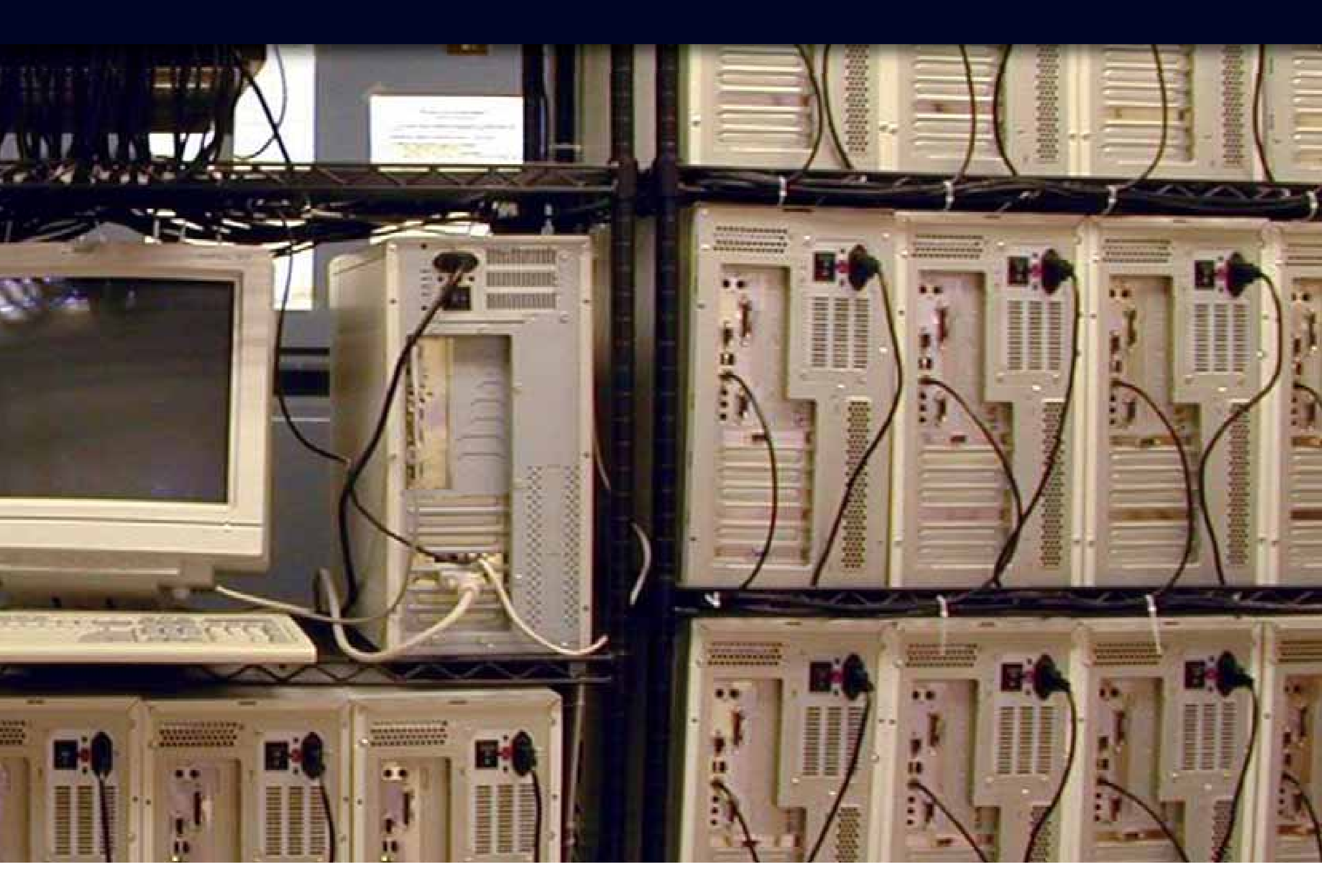
Autor: **Reinaldo Pis Diez**

A más de diez años del nacimiento del cluster Beowulf, muchísimos equipos del mundo son clusters que corren bajo sistema operativo Linux. Pero, ¿qué es un cluster Beowulf? ¿Cómo se instala? En este artículo vamos a tratar estos temas e intentar ir un poco más allá.

Un poco de Historia

La noción de aunar el poder individual de las computadoras para lograr una mayor potencia de cálculo no es novedosa en absoluto. En las décadas del '50 y '60, la Fuerza Aérea norteamericana construyó una red de computadoras (funcionando con válvulas en lugar de circuitos integrados) como método de defensa contra un eventual ataque nuclear de la entonces Unión Soviética. Denominaron a esa red SAGE [1]. A mediados de los '80, la Digital Equipment Corporation acuñó el término "cluster" para bautizar su sistema de minicomputadoras VAX [1].

A comienzos de los '90, tanto la caída en los precios de las computadoras personales, mejor conocidas como "PCs", como el rápido desarrollo de la tecnología Ethernet, dominante para conectar computadoras en una red local, condujeron de manera lógica a la idea de utilizar PCs como "unidades de construcción" de un cluster. Sin embargo, existía una dificultad importante: el sistema operativo de las PCs no era tan flexible ni tan poderoso como el sistema operativo UNIX, utilizado en la mayoría de las computadoras de mayor porte. Hubo que esperar muy poco tiempo hasta que Linus Torvald presentara en sociedad y ofreciera gratuitamente su sistema operativo "Linux", un UNIX especialmente diseñado para instalar en PCs. Estaban dadas las condi-



ciones para construir un cluster de PCs con sistema operativo Linux [1].
En 1994, el Centro Espacial Goddard de la NASA presentó un cluster de 16 PCs, todas ellas con procesador Intel486, utilizando Linux como sistema operativo y Fast Ethernet como tecnología de conexión. Obtuvieron una potencia máxima de cálculo de 70 megaflops (1 megaflops = 1 millón de operaciones de punto flotante por segundo), similar a la brindada por sistemas comerciales cuyo valor era de 10 veces superior al del cluster de PCs. Denominaron al cluster "Beowulf", en honor al rey anglosajón del medioevo que derrotó al monstruo Grendel arrancándole un brazo [1].

A más de diez años del nacimiento del cluster Beowulf, la página Top500.org permite verificar que 367 equipos de los 500 más veloces del mundo son clusters que corren bajo sistema operativo Linux [2], con rendimientos que alcanzan los 10 TFlops.

¿Qué es "High Performance Computing"?

El término "High Performance Computing" (computación de alto rendimiento) tiene su origen en el trabajo desarrollado por Seymour Cray al diseñar y construir las computadoras que llevan su nombre. Esos sistemas vectoriales proporcionaban un poder de cómputo, un orden de magnitud mayor que otros sistemas contemporáneos, razón por la cual fueron denominadas "supercomputadoras".
El rápido incremento mostrado en nuestros días tanto por la velocidad de los procesadores de computadoras personales como por la relación rendimiento/precio que las mismas suministran, sumado al desarrollo de tecnologías de redes de alta velocidad, conducen necesariamente a una profunda transformación del concepto original de High Performance Computing.
Actualmente, pensar en un sistema para efectuar computación de alto rendimiento es pensar en un cluster de computadoras personales utilizando el sistema operativo Linux y efectuando sus tareas bajo un entorno de programación en paralelo.
Muchas son las disciplinas que requieren llevar a cabo diariamente computación de alto rendimiento como una parte importante de sus actividades. Algunas de ellas son la astronomía, la biología, la física, muchas ramas de la ingeniería, y la química. El procesamiento de imágenes, la determinación de la secuencia de los nucleótidos en el ADN y la simulación numérica de procesos de interés biológico, químico o físico son algunas de las actividades que requieren de la computación de alto rendimiento para alcanzar resultados en tiempos razonables. Podemos decir entonces que la "computación de alto rendimiento" es aquella dedicada a la resolución de problemas bien determinados por medio de sistemas de altísimo poder de cómputo.

¿Qué es un cluster Beowulf?

No existe una única definición de cluster Beowulf, pero si nos atenemos a la configuración original el mismo debería consistir en una PC cumpliendo el rol de servidor y una o más PCs clientes conectadas a aquel por una red Fast Ethernet o mejor. Debe construirse utilizando hardware que pueda adquirirse en cualquier negocio especializado con el fin de que las dificultades o desperfectos puedan subsanarse inmediatamente y, además, su configuración pueda reproducirse con facilidad. El software debe también cumplir con el requisito de gratuidad o, eventualmente, bajo costo.
El servidor administra los clientes, las tareas asignadas en éstos y es además la conexión con el mundo exterior. Es deseable que los clientes sean máquinas "bobas" en el sentido que ni siquiera presenten una pantalla para ingresar nombre de usuario y clave. En este sentido, un cluster Beowulf puede pensarse como una única máquina con unidades de CPU y memoria que pueden darse de alta o de baja en cualquier momento. Esta es una diferencia importante con un cluster de esta-

ciones de trabajo, donde cada uno de sus componentes es en sí mismo una máquina independiente de la cual, en ciertas ocasiones, se requiere su servicio para ejecutar aplicaciones en paralelo.

Hardware e Instalación de un Cluster Beowulf

La elección del hardware para construir el cluster Beowulf depende fuertemente de las aplicaciones a ejecutar en el mismo. Es importante conocer a priori el factor limitante de dichas aplicaciones. Una aplicación que se ejecuta en paralelo puede estar limitada por el poder de cómputo o bien por la cantidad de lectura/escritura de datos en los discos duros. Así, una aplicación limitada por el poder de cómputo requerirá de veloces procesadores y una red de comunicaciones de alta velocidad para alcanzar la eficiencia deseada. Por otro lado, una aplicación limitada por lectura/escritura aumentaría su eficacia en un sistema con procesadores de mediana a baja velocidad y una red de comunicaciones tipo Fast Ethernet o GigaNet.
La instalación de un cluster Beowulf no es una tarea complicada pero requiere atender especialmente algunos detalles. Es conveniente que el cluster constituya una red privada, es decir, que utilice direcciones de tipo 192.168.x.x ó 10.0.x.x. Si el servidor acepta conexiones desde otra red interna o desde el exterior, el mismo deberá poseer dos placas de red. La instalación del sistema operativo en el servidor debe incluir el software NFS (Network File System). Toda distribución Linux actual incluye ese software en su paquete base. De esta forma, las cuentas de usuarios y el software y bibliotecas necesarias para ejecutar aplicaciones en paralelo residen únicamente en el servidor y los clientes acceden a los mismos montando el sistema de archivos o directorio correspondiente. Así, las tareas de administración, mantenimiento y actualización se ven notoriamente facilitadas.
Siguiendo la misma filosofía, es altamente conveniente que los clientes sean exactamente idénticos desde el punto de vista de las características del sistema operativo instalado.
Debido a la necesidad de que los paquetes de información migren libremente entre los nodos, el servidor de SecureShell debe estar activo en el servidor y los nodos, y por supuesto deben existir los clientes en cada máquina. Si el cluster se encuentra aislado de internet, los comandos rsh, rlogin y rexec pueden utilizarse como una alternativa al SecureShell.
Finalmente, las bibliotecas PVM (Parallel Virtual Machine) [4] y MPI (Message Passing Interface) [5], esta última en sus versiones LAM-MPI [6] o MPICH [7], son el único software imprescindible para poder desarrollar y ejecutar aplicaciones en un cluster Beowulf.

Hacia “Beowulf 2”

Muy recientemente se ha introducido el concepto de “Beowulf 2” para indicar importantes modificaciones en lo que hace a la instalación, administración y distribución de procesos en un cluster Beowulf [8, 9]. Esas modificaciones parten de reconocer algunos defectos de los cluster Beowulf originales: adicionar nuevos clientes implica instalar el sistema operativo en ellos y modificar los archivos necesarios en el servidor y en los viejos clientes para que los nuevos sean reconocidos; una actualización de kernel importante hace necesaria la instalación de los parches y la recopilación de aquel en todos los nodos del cluster; etc.
La aparición de paquetes como Scyld [10], OSCAR [11] o RocksClusters [12] tiene la intención de facilitar no sólo la instalación y administración del cluster, sino también las tareas de adicionar nuevos clientes y actualizar versiones de kernel. A modo de ejemplo describiremos con algún detalle la instalación de un cluster Beowulf utilizando el paquete Scyld.

Instalando un cluster “Beowulf 2” con Scyld

La versión básica, sin soporte ni documentación, del paquete Scyld puede ser descargada desde ftp://ftp.scyld.com/pub. La última versión disponible en ese sitio es la 29cz. Como se comentó anteriormente, es recomendable que la configuración del cluster sea como la que se muestra en la figura 1, ya que la instalación permite ajustar los parámetros de dos tarjetas de red.
No es necesario que el servidor tenga una distribución de Linux preinstalada ya que Scyld trae la versión 4 de CentOS, una excelente altenativa gratuita a Red Hat Enterprise Server. Una vez establecido en el BIOS el inicio desde la lectora de cd, lo primero que se observa es una instalación típica de CentOS. Es altamente recomendable elegir el modo gráfico de instalación y luego una instalación completa con soporte gráfico en el escritorio

¿Necesito un cluster Beowulf para mis actividades?

No todas las actividades precisan de un alto poder de cómputo para su realización. Por ejemplo, construir un cluster de 4 procesadores para utilizar el procesador de textos favorito es verdaderamente un desperdicio. Sin embargo el manejo de grandes bases de datos, el tratamiento de imágenes satelitales o la resolución de problemas numéricos en diversos campos de la ciencia justifica plenamente la construcción o adquisición de tal cluster o de uno más poderoso. Básicamente uno podría preguntarse:
• ¿Tengo limitaciones de velocidad, memoria, etc. para llevar a cabo mis tareas?
• ¿Tengo programas o paquetes de programas que permiten ser ejecutados en paralelo utilizando el paradigma de “message passing”?
• ¿Los programas que desarrolla mi grupo pueden ser adaptados para ser ejecutados en paralelo bajo el paradigma anterior?
Si las respuestas a estas preguntas son positivas, entonces debería plantearse seriamente la posibilidad de asesorarse y, eventualmente, construir o adquirir un cluster Beowulf.
Si bien es cierto que todas las ventajas de un cluster Beowulf pueden ser obtenidas por sistemas comercializados por empresas reconocidas internacionalmente, es mundialmente aceptado que un sistema Beowulf proporciona el mismo rendimiento a un costo por lo menos diez veces menor al del sistema comercial.

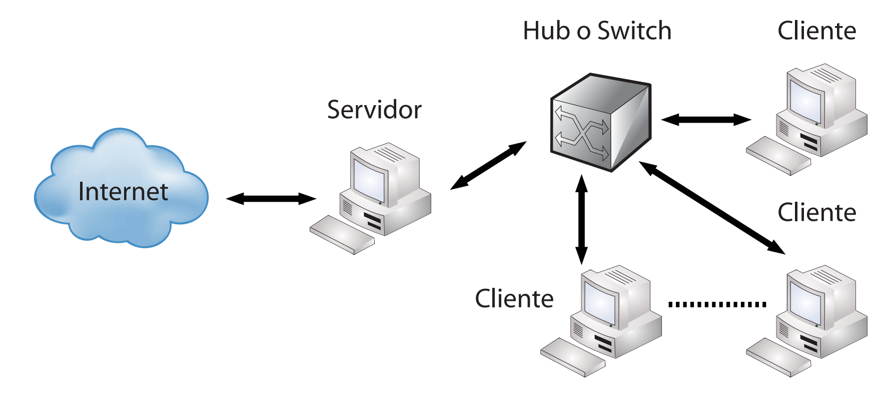


Fig.1 Cluster “Beowulf 2” con Scyld.

RocksClusters

ROCKS

RocksClusters es una distribución Linux para instalar clusters de tipo HPC desarrollada en el Centro de Supercomputación San Diego de la Universidad de California en San Diego, UCSD. Una de las particularidades de esa distribución es que permite diferenciar distintos roles a cumplir por los componentes del cluster: desde servidores de archivos hasta nodos de cómputo, pasando por nodos de monitoreo. La instalación de un cluster HPC completo es altamente automatizada y se realiza, literalmente, en pocas horas. Está basada en la distribución CentOS e incluye distintos "rolls" que pueden instalarse según las necesidades, como por ejemplo, el administrador Ganglia, los administradores de colas PBS/Maui y Sun Grig Engine (SGE) o bien el basado en el software Globus para convertir a nuestro cluster en parte de un sistema GRID. Puede aprender más de este producto leyendo la documentación en http://www.rocksclusters.org.

Gnome, ya que algunas herramientas del paquete están sólo disponibles en ese modo. Si se tiene suficiente experiencia en instalación de Linux, se puede elegir el modo "custom" ya que entonces se evita la instalación de muchas aplicaciones que casi seguramente nunca se usarán. En este caso, es conveniente permitir la instalación del modo gráfico con el escritorio Gnome por las razones arriba mencionadas.

Solo una diferencia se puede apreciar respecto a una instalación típica de RedHat: una ventana dedicada a configurar la interfaz eth1, la segunda tarjeta de red. Si se escoge, por ejemplo, una dirección privada 192.168.1.1 para el servidor, entonces el rango de direcciones a usar debe comenzar en 192.168.1.2 y terminar en 192.168.1.99, siendo 99 un número tentativo (¡y optimista!), por supuesto.

Una vez finalizada la instalación del sistema operativo, el servidor se reinicia y, luego de cargar el escritorio automáticamente, se inician tres ventanas: una guía de instalación rápida de los clientes, la aplicación BeoSetup para instalarlos y la aplicación BeoStatus para monitorear el estado del cluster.

Los clientes deben poder iniciar desde el cd de Scyld o desde un disquete que haya sido generado con la utilidad BeoBoot. Ya que resulta engorroso estar llevando el cd de un cliente a otro, es conveniente generar un disquete por cada cliente del cluster. Para tal fin se encuentra un botón en la ventana de BeoSetup del servidor. Una vez generados los disquetes, se usan para iniciar los clientes.

Lo que sigue es muy interesante: cada cliente envía al servidor un requerimiento de dirección RARP (Reverse Address Resolution Protocol). El servidor acepta los requerimientos y las direcciones físicas de aquellos que aparecen en la ventana derecha de la aplicación BeoSetup, "Unknown Addresses". Dichas direcciones se arrastran con el ratón a la ventana central de BeoSetup, "Configured Nodes", y, automáticamente, el servidor envía una orden de reinicio a cada cliente. El cambio de estado de los clientes resulta evidente al ver la palabra "up" en la columna "Node Status" de la ventana central de BeoSetup. Al mismo tiempo, BeoStatus refleja los cambios producidos al dar de alta los clientes en el cluster.

Los discos duros de los clientes se particionan con la utilidad beofdisk que debe ser invocada desde una consola:

#>beofdisk -d

#>beofdisk -w

La opción *-d* usa la configuración predeterminada para particionar los discos de todos los clientes y *-w* ocasiona la escritura de la tabla de partición en los mismos. Para que estos cambios surtan efecto, deben activarse las líneas con información de los sistemas de archivos */dev/hda2* y */dev/hda3* en el archivo */etc/beowulf/fstab* y comentar la línea que comienza con $RAMDISK. Si se desea transferir la imagen de inicio de BeoBoot al disco duro de los clientes, debe ejecutarse:

#>beoboot-install -a /dev/hda

El último paso implica reiniciar todos los clientes, lo cual se efectúa por medio de un único comando desde el servidor

#>bpctl -S all -s reboot

De esta forma, se tiene un cluster Beowulf instalado y operativo en menos de una hora. Como broche de oro, el paquete Scyld permite comprobar el rendimiento del cluster por medio de la utilidad Linpack [13]. ●



Referencias

[1] William W. Hargrove, Forrest M. Hoffman, Thomas Sterling, The Do-It-Yourself Supercomputer, edición digital de Scientific American (www.scientificamerican.com), agosto 2001.
[2] www.top500.org
[3] www.oracle.com/ip/deploy/ias/Linux
[4] www.epm.ornl.gov/pvm/
[5] www.mcs.anl.gov/mpi/index.html
[6] www.lam-mpi.org
[7] www-unix.mcs.anl.gov/mpi/mpich
[8] ww.linuxjournal.com/article.php?sid=5559
[9] www.linuxjournal.com/article.php?sid=6019
[10] www.scyld.com
[11] http://oscar.sourceforge.net, www.linux-journal.com/print.php?sid=5559
[12] www.rocksclusters.org/
[13] www.netlib.org/benchmark/hpl/



Overview del
Windows Compute Cluster Server 2003

Por: **Charlie Russel**
Microsoft MVP para Windows Server

Este artículo da una visión general del Microsoft Windows Compute Cluster Server 2003 y detalla sus requerimientos, tecnologías y beneficios. Se describe la arquitectura del sistema junto con los componentes clave y los beneficios del Windows Compute Cluster Server 2003, incluyendo la administración e implementación, y el Microsoft Message Passing Interface (MS-MPI), programador de tareas y seguridad.



¿Qué es Windows Compute Cluster Server 2003?

Microsoft ha participado del mundo de clusters pero restringido a la idea de tolerantes a fallas (fault-tolerant) y de alta disponibilidad (high-availability) por muchos años, pero con el lanzamiento de Windows Compute Cluster Server (WCCS) 2003, Microsoft logra llevar el poder de súper-cómputo de la computación de alta performance (HPC: High Performance Computing) al nivel personal y de grupo de trabajo.

Historia

En 1991, una supercomputadora de 10-Gflops (1 FLOPS = 1 Floating point Operation Per Second) era una Cray que costaba aproximadamente 40 millones de Dólares. Hoy, ese mismo poder de cómputo puede obtenerse combinando cuatro computadoras x64 a un costo de alrededor U$S 4.000, logrando así bajar el costo de adquisición del hardware de la supercomputación al nivel del de escritorios personales.
Sin embargo, un gran reto para hacer realidad la supercomputación personal ha sido la complejidad de su implementación y administración. Windows Compute Cluster Server 2003 integra la implementación de clusters y la administración en las familiares herramientas del servicio de directorio de Microsoft Active Directory y Microsoft Windows Server™ 2003, permitiendo, mediante el uso de Microsoft Management Console (MMC), una implementación simple de computadoras nodos y una administración centralizada.

¿Cómo se ofrece WCCS 2003?

Windows Compute Cluster Server 2003 es un paquete de dos CDs: el primero contiene Windows Server 2003 Compute Cluster Edition y el segundo, el Microsoft Compute Cluster Pack, una combinación de interfaces, utilidades e infraestructura de administración que constituyen Windows Compute Cluster Server 2003.

Requerimientos del Hardware

Los requerimientos mínimos del sistema (ver Tabla 1) son similares a los necesarios para Windows Server 2003 Standard x64 Edition. Windows Compute Cluster Server 2003 soporta hasta 32 gigabytes (GB) de RAM.
Los microprocesadores soportados son:
- AMD Opteron
- AMD Athlon 64
- Intel Xeon con Intel EM64T
- Intel Pentium con Intel EM64T

Requerimientos del Software

Para WCCS 2003 el nodo principal y los de cómputo pueden ser cualquiera de los siguientes sistemas operativos:
- Windows Server 2003 Compute Cluster Edition
- Windows Server 2003 Standard x64 Edition
- Windows Server 2003 Enterprise x64 Edition
- Ediciones x64 de Windows Server 2003 R2

Los componentes de administración remota y programación de trabajos son instalados automáticamente en el nodo principal del cluster de cómputo pero pueden ser también instalados en una estación de trabajo remota para simplificar la administración y la programación.
Los sistemas operativos soportados para la instalación de los componentes remotos son:
- Windows Server 2003 Service Pack 1 ó R2
- Microsoft Windows® XP Professional Service Pack 2.
- Windows XP Professional x64 Edition

Nota: Windows Server 2003 Compute Cluster Edition está diseñado solamente para uso con aplicaciones de alta performance y no soporta ser usado como un servidor para una infraestructura de uso general.

Arquitectura de Solución

WCCS 2003 es un cluster de servidores que incluye un único nodo principal y uno o más nodos de cómputo (Figura 1). El nodo principal controla y media entre todos los accesos a los recursos del cluster. WCCS 2003 utiliza la existente infraestructura corporativa de Active Directory para operaciones de seguridad, administración de cuentas y para la administración de operaciones en general, usando herramientas como Microsoft Operations Manager 2005 y Microsoft Systems Management Server 2003.
La instalación del WCCS 2003 concierne en instalar el sistema operativo en el nodo principal, unirlo a un dominio Active Directory



Tabla1: Requerimientos Mínimos del Hardware

CPU	Arquitectura de x64 con procesador de la familia Intel Pentium o Xeon con arquitectura de procesador Intel Extended Memory 64 Technology (EM64T); familia de procesadores AMD Opteron; familia de procesadores AMD Athlon; procesadores compatibles.
RAM	512 MB
Soporte Multiprocesador	Windows Compute Cluster Server 2003 y Windows Server 2003 Standard x64 Edition soportan hasta cuatro procesadores por servidor. Windows Server 2003 Enterprise x64 Edition soporta hasta ocho procesadores por servidor.
Espacio en disco para la instalación	4 GB
Unidades de Disco	Si el Remote Installation Service (RIS) es usado, se requieren dos unidades (C:\ y D:\) en el nodo principal; una para la partición del sistema y la otra para el RIS. Si el nodo principal está siendo usado para tareas adicionales, como ser un servidor de archivos o de impresión, se recomiendan particiones adicionales. Se requiere de un sistema con una sola unidad para los nodos de cómputo. Los discos RAID (Redundant array of independent disks) son soportados pero no requeridos.
Network Interface card	Se requiere por lo menos una placa de red (NIC, Network Interface Card). Si se usa red privada, el nodo principal necesita por lo menos dos NICs, y los nodos de cómputo por lo menos una. Cada nodo podría también necesitar una NIC de alta velocidad para una red Message Passing Interface (MPI).

existente, y luego instalar el Compute Cluster Pack. Si Ud. llegara a estar usando RIS para implementar automáticamente nodos de cómputo, RIS se instalará y configurará como parte del To Do List (Lista de tareas pendientes) luego de que la instalación se complete.

Cuando la instalación del Compute Cluster Pack se haya completado, mostrará una página del To Do List que muestra los pasos necesarios para completar la configuración de su cluster de cómputo. Estos pasos incluyen definir una topología, configurar RIS utilizando el asistente Configure RIS, agregar nodos de cómputo al cluster y configurar usuarios y administradores del cluster.

Topología de Red

Windows Compute Cluster Server 2003 soporta cinco diferentes topologías de red, con una y hasta tres NICs en cada nodo. Las cinco topologías soportadas son:

• *Tres NICs en cada nodo.* Una NIC es conectada a la red pública (corporativa); una a la red de administración privada del cluster; y otra a una red MPI (Message Passing Interface) de alta velocidad.

• *Tres NICs en el nodo principal y dos en cada uno de los nodos del cluster.* El nodo principal provee NAT (Network Address Translation) entre los nodos de cómputo y la red pública, con cada nodo de cómputo teniendo una conexión a la red privada y una conexión a un protocolo de alta velocidad como MPI.

• *Dos NICs en cada nodo.* Una NIC es conectada a la red pública (corporativa), y una es conectada a la red cluster privada y dedicada.

• *Dos NICs en el nodo principal, y uno en cada uno de los nodos de cómputo.* El nodo principal provee NAT entre los nodos de cómputo y la red pública.

• *Una solo NIC en cada nodo, con todo el tráfico de red compartiendo la red pública.* En este limitado escenario de red, la implementación RIS de nodos de cómputo no está soportado, y cada nodo de cómputo debe ser instalado y activado manualmente.

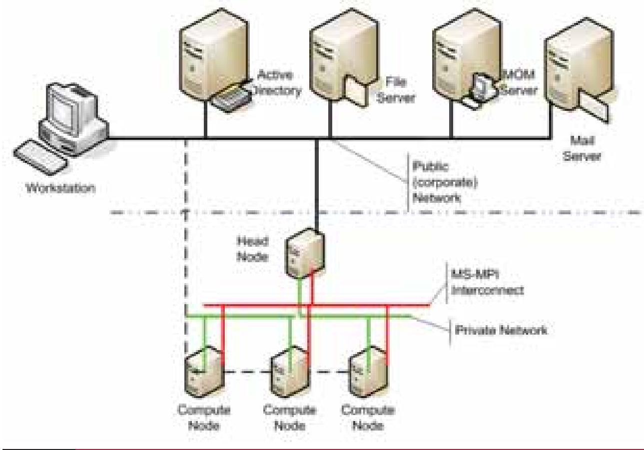


Fig. 1 Típica red Windows Compute Cluster Server 2003

La Microsoft Message Passing Interface (MS-MPI) es una red de alta velocidad que corre sobre Gigabit Ethernet, InfiniBand, o cualquier otra red que provea un driver que permita WinSock Direct. MSMPI está basado y es compatible con la implementación Argonne National Labs MPICH2 de MPI2.

Beneficios y Elementos de WCCS 2003

WCCS 2003 provee una ponderosa solución HPC de bajo costo que corre en computadoras x64 y puede ser fácil y rápidamente expandido e implementado usando tecnologías y herramientas familiares.

Administración e Implementación

Uno de los grandes problemas que afrontan hoy los clientes al adoptar soluciones HPC es la administración e implementación de clusters y nodos. Este problema ha sido tradicionalmente un problema a nivel departamental o corporativo, con un equipo de profesionales de TI dedicado a manejar e implementar nodos, y usuarios enviando lotes de trabajo y compitiendo por recursos limitados. Las metas del diseño de WCCS 2003 son:

• Proveer una configuración similar a un *appliance*.

• Dar una guía clara y precisa.

• Proveer mecanismos de autenticación y autorización.

• Crear una solución [*scriptable*].

WCCS 2003 potencia Active Directory y

MMC 3.0 permitiéndole proveer una interfaz simple y familiar para manejar y administrar el cluster. Integrándolo con el Active Directory permite una fácil administración del cluster basada en los roles. El nuevo Compute Cluster Administrator tiene cinco páginas importantes:

• ***Página de Inicio.*** Es primariamente una página de monitoreo. Esta página muestra el número de nodos y sus estatus, el número de procesadores en uso y disponibles, e información de trabajos y su estatus.

• ***Lista de Tareas Pendientes (To Do List).*** Esta página es usada para configurar y administrar el cluster, incluyendo la red, RIS, agregar o quitar nodos y seguridad del cluster.

• ***Administración de Nodos.*** Esta página expone información acerca de nodos y trabajos en el cluster y permite diferentes tareas en los nodos, como ser aprobar un nodo, pausar o continuar un nodo o resetearlo.

• ***Sesiones de Escritorio Remoto*** Esta página es usada para crear y cerrar sesiones de escritorio remotas en los nodos de cómputo.

• ***Monitor de Performance.*** Esta página muestra datos del monitoreo de la performance de PerfMon, incluyendo tiempo y trabajos de procesador y estadísticas por cada nodo.

A parte del Compute Cluster Administrator, existe un Compute Cluster Manager que es usado para envíos y administración de trabajos, y un Command Line Interface (CLI) que provee una alternativa a la línea de comandos para administrar el cluster y manejar los trabajos.

La configuración e implementación son simplificadas en gran medida con Windows Compute Cluster Server 2003. La instalación inicial del nodo principal saca ventaja de los asistentes para simplificar e identificar los pasos necesarios, mientras que el uso de RIS hace agregar un nodo de cómputo.

MPI

La Microsoft Message Passing Interface (MS-MPI) es una versión de la implementación Argonne National Labs Open Source MPI2 la cual es ampliamente usada por clusters HPC existentes. MS-MPI es compatible con MPICH2 Reference Implementation y con otras implementaciones MPI y soporta una API con más de 160 llamadas de función.

La MS-MPI en WCCS 2003 mejora el protocolo WinSock Direct para una mejor performance y eficiencia del CPU. MS-MPI puede utilizar cualquier interconexión Ethernet que sea soportada por Windows Server 2003 así como interconexiones de baja latencia y de banda ancha, como por ejemplo InfiniBand o Myrinet, a través de los drivers Winsock Direct provistos por los fabricantes de hardware. Gigabit Ethernet brinda una tela de interconexión de alta velocidad y a bajo costo, mientras que InfiniBand es ideal para aplicaciones de banda ancha y sensible a la latencia. MS-MPI incluye soporte para los lenguajes de programación C, Fortran77, y Fortran90; y la última versión del Microsoft Visual Studio incluye un *debugger* paralelo que funciona con MS-MPI. Los desarrolladores pueden ejecutar sus aplicaciones MPI en múltiples nodos de cómputo desde dentro del *ambiente* Visual Studio, y luego Visual Studio conectará automáticamente los procesos en cada nodo, permitiendo al desarrollador pausar individualmente y examinar las variables del programa en cada nodo.

¿Qué es el FLOPS?

El término Floating point Operations Per Second (operaciones de punto flotante por segundo) se usa como una medida del rendimiento de una computadora, especialmente en cálculos científicos que requieren un gran uso de operaciones de punto flotante.
Para marcar la diferencia una computadora moderna de escritorio, que usa por ejemplo un procesador Pentium 4 o Athlon 64, típicamente opera a más de 2 GHz y provee de un desempeño computacional del rango de unos cuantos GFLOPS (10^9 FLOPS). Sin embargo, una de las computadoras más rápida, la supercomputadora IBM Gene Blue/L, es capaz de llegar a un pico máximo de 280,6 TFLOPS (10^{12} FLOPS).

Programador

WCCS 2003 incluye un programador de tareas de línea de comandos; y el Compute Cluster Manager que le permite a los usuarios programar trabajos, ubicar los recursos necesarios para el trabajo, y cambiar las tareas y propiedades asociadas con éste.

El CLI soporta varios lenguajes, entre ellos Perl, Fortran, C/C++, C#, y Java. Las tareas pueden ser únicas o múltiples y especificar el número de procesadores requeridos para el trabajo y si esos procesadores son necesitados exclusivamente o pueden ser compartidos con otros trabajos o tareas.

Las características más importantes que distinguen al Programador son:

• ***Recuperación de Errores.*** Permite el reintento automático de tareas y trabajos fallidos, y routeo automático de los nodos que no responden. También provee la detección de aquellos nodos que no dan respuesta.

• ***Limpieza Automatizada.*** Cada proceso asociado a un trabajo o tarea es monitoreado y proactivamente apagado en todos los nodos de cómputo al concluir la tarea o trabajo, previniendo que se ejecuten procesos en los nodos de cómputo.

• ***Seguridad.*** Cada trabajo o tarea corre en el contexto del usuario que lo haya enviado y mantiene la seguridad a lo largo de todo el proceso.

Seguridad

Siendo los clusters HPC la elección de una amplia gama de usuarios para las aplicaciones críticas, la seguridad e integración con la infraestructura existente son esenciales. WCCS 2003 le permite al Active Directory habilitar la seguridad basada en roles para todos los trabajos de los clusters y administración. El programador corre cada trabajo bajo el contexto y las credenciales del usuario que lo remite, no las de un súper usuario, y todas las credenciales son almacenadas con el trabajo y eliminadas cuando éste concluye. Este comportamiento les permite a los nodos de cómputo acceder a los recursos de red, como servidores de archivos o bases de datos, en el contexto del usuario y permite a los administradores de sistema aplicar y auditar políticas de seguridad usando los mecanismos existentes y familiares en Active Directory.

Todas las comunicaciones de administración de los trabajos son realizadas sobre canales encriptados y autenticados, y las credenciales son sólo conocidas por el administrador de nodos durante el tiempo que toma el trabajo. El proceso de cómputo en sí mismo sólo ve un logon token, no las credenciales actuales, aislándolas y protegiendo su integridad.

WCCS 2003 ayuda a proveer seguridad *end-to-end* sobre canales seguros y encriptados a lo largo del proceso del trabajo cuando se utiliza MS-MPI. Mientras que el administrador de nodos programa y asigna el trabajo, y las tareas son distribuidas, el trabajo siempre corre en el contexto del usuario del programador. Éste es un importante agregado a la implementación MS-MPI que no forma parte de la implementación de referencia MPICH2.

Conclusión

WCCS 2003 reúne el poder de las computadoras x64, la facilidad de uso y seguridad del Active Directory, y el sistema operativo Windows para dar una solución HPC segura y costeable. WCCS 2003 puede ser fácil y rápidamente implementado usando tecnologías estándar de implementación de Windows, y nodos de cómputo adicionales pueden ser agregados al cluster de cómputo simplemente enchufando los nodos y conectándolos. La Microsoft Message Passing Interface es totalmente compatible con la referencia MPICH2. La integración con Active Directory ayuda a permitir la seguridad basada en roles para administradores y usuarios, y el uso de Microsoft Management Console brinda una interfaz familiar de administración y programación de trabajos. ●

Computación de Alto Rendimiento
Clusters y Grid

Cómo se aplica HPC en el marco de la investigación universitaria y qué proyectos se llevan adelante dentro de una red informática que conecta a Buenos Aires y Mendoza.

Hoy en día se necesitan computadoras muy poderosas y costosas para resolver ciertos cálculos complejos, entre ellos tareas relacionadas con las finanzas, simulaciones meteorológicas y astronómicas, computación gráfica, como así también en el área de la tecnología biomédica entre otros. En esta nota nos centraremos en los diversos proyectos que necesitan "High Performance Computing (HPC)" dentro del marco de la investigación. Más precisamente en los que se llevan adelante en el Laboratorio de Sistemas Complejos (LSC) del Departamento de Computación de la Facultad de Ciencias Exactas y Naturales (FCEN) de la UBA con el Laboratorio de Investigación y Desarrollo para la Producción Integrada por Computadora (LAPIC) de la Universidad Nacional de Cuyo (UNCU) en Mendoza.

Dichos proyectos utilizan dos clusters de alto rendimiento, pertenecientes a los laboratorios de LSC y UNCU, conectados entre sí a través de la red académica. De esta forma se comparte información y se intensifica el poder de cálculo de los dos laboratorios en forma virtual. Este centro de Computación de Alto Rendimiento, que lleva tres años en producción, fue posible gracias a un subsidio de investigación para un proyecto en red otorgado por la ANPCyT, proyecto que dirigen Guillermo Marshall en la Universidad de Buenos Aires y Carlos García Garino en la Universidad de Cuyo, Mendoza.



Cómo funciona este Grid

El término Grid fue tomado de la expresión "power grid", que es como se llama a las redes de energía eléctrica en los países de habla sajona. La idea que subyace a esta analogía supone la posibilidad de conectar diferentes recursos (clusters, workstations, PCs individuales) y obtener un alto poder de procesamiento sin importar qué componentes del grid proveen la capacidad de cómputo, de la misma manera que se obtiene electricidad cuando se conecta un artefacto a la red eléctrica sin interesar dónde se generó esa energía.

Grid ha recibido un considerable impulso en los países de economías desarrolladas (Estados Unidos, Canadá, Reino Unido, Australia) e incluso en algunos países de América Latina como Brasil, pero es relativamente poco conocida en nuestro país. La puesta en marcha de un entorno Grid para permitir que se compartan de manera segura y eficaz los recursos a disposición de la red académica es una tarea de gran envergadura. En particular este Grid universitario se compone actualmente de un cluster en Buenos Aires en el LSC y de otro en Mendoza en el

Beneficios del Grid

La utilización de una infraestructura Grid permite a los investigadores:

- **Intercambiar conocimientos**
- **Disminuir los costos de equipamiento al compartir recursos existentes. No es necesaria la aplicación de anexos con altos presupuestos.**
- **Aumentar la capacidad de cómputo. El funcionamiento de todos los clusters al mismo tiempo logran mayor rapidez para obtener resultados.**
- **Optimizar el uso de recursos.**

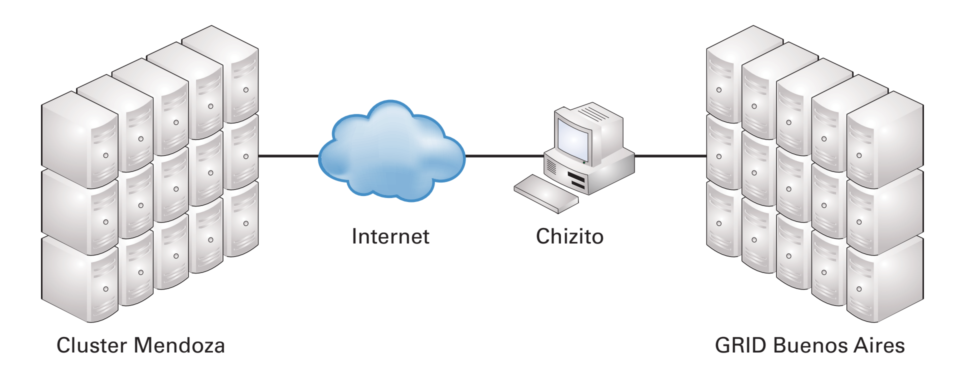


Fig.1 Interconexión de nodos Grid entre el LSC y el cluster ubicado en el LAPIC

LAPIC. Grid Computing permite por ejemplo en este caso, la integración de recursos informáticos administrados por diversas organizaciones geográficamente alejadas, asumiendo la ausencia de una ubicación central de control. Para que la relación entre los dos nodos principales que conforman este Grid pueda ser establecida de forma correcta, es necesaria la participación de un middleware. Es decir, una infraestructura virtual o capa de software que media entre las computadoras para así dividir el trabajo más eficientemente, compartir el costo de infraestructura y aumentar el poder de cada máquina por permanecer conectada con las demás.

En una red Grid una computadora puede solicitar algún servicio (poder de cómputo para una tarea específica) a las otras sin importar que sus hardwares y softwares sean diferentes. Para lograrlo, el middleware determina qué computadora está disponible y qué facilidades posee para proveer el servicio requerido. Luego de verificar la capacidad ociosa de cada integrante de la red, puede asignar las porciones de trabajo.

Para que los diferentes procesadores que conforman un Grid logren los objetivos específicos que se les asignan deben funcionar de forma coordinada y como una única máquina. Los componentes básicos que conforman este Grid son clusters y en ellos se realiza lo que en la jerga informática se llama "paralelización". Si uno imagina un programa de computación clásico (secuencial) como un conjunto de instrucciones operando sobre un conjunto de datos, una manera sencilla de paralelizar esta operación consiste en distribuir los datos entre todas las computadoras, de modo que cada una de ellas ejecute las mismas instrucciones en su porción de los datos.

Investigaciones

En el LSC se realizan diversos proyectos relacionados a Grid Computing basados en herramientas Open Source. Uno de los principales es, mediante la creación de dispositivos móviles en conjunción con sistemas portátiles de adquisición de datos (tales como electrocardiógrafos o monitores de presión arterial), la confección de una base de datos clínicos (enfermedades cardiovasculares, diabetes y cáncer) para su uso en medicina y para el procesamiento en centros médicos especializados para el posterior tratamiento del enfermo.

Con el objetivo de promover la investigación y desarrollo en América Latina, el proyecto de Guillermo Marshall y Marcelo Risk, miembros del CONICET y Profesores de la UBA que se realiza en el LSC, Departamento de Computación, FCEyN, UBA, propone la construcción de un kit de inclusión digital para la expansión de las fronteras de la tecnología de la computación en la salud. Asimismo, con esta información se creará una base general de datos clínicos para el análisis de la información y la aplicación de medicina preventiva mediante el uso de un Cluster Virtual (VC) que permite compartir la información entre comunidades educativas y de investigación. Este proyecto esta financiado por Microsoft Research RFP2006 Digital Inclusion.

Sin embargo, las investigaciones y los avances logrados están abocados al estudio de los sistemas complejos en bioelectroquímica hidrodinámica, al tratamiento electroquímico de tumores de ratón, in vivo e in vitro, y a la simulación numérica en máquinas secuenciales y en paralelo en contextos de HPC y Grid. La coordinación de este proyecto está a cargo de Guillermo Marshall junto con un grupo interdisciplinario formado con investigadores del Departamento de Computación, del Departamento de Química General Inorgánica y Química Física (INQUIMAE), FCEyN, UBA, y con la colaboración del Dr. Lucas Colombo, Investigador del Conicet en el Instituto Roffo de Oncología de la Universidad de Buenos Aires. En el tratamiento electroquímico de tumores (EChT) el tejido tumoral es tratado con una corriente continua a través de dos o más electrodos ubicados cerca o en el tumor. Sus principales ventajas son la seguridad, menor trauma y menor tiempo de recuperación del paciente. El EChT es estudiado en China, país donde se realizaron ensayos clínicos muy promisorios (más de diez mil pacientes tratados con EChT durante los últimos diez años). El objetivo final del estudio de EChT es demostrar cómo, a través de la simulación numérica de alto rendimiento, esta puede ser utilizada como una herramienta valiosa para definir y optimizar las condiciones de operación del tratamiento electroquímico de tumores. ●



Cluster Beowulf del LSC del Departamento de Computación en FCEyN



Parte del grupo de trabajo del LSC; a la derecha Guillermo Marshall

Soluciones libres al problema del load-balancing en clusters

openMosix y Condor

Por: **Dr. Reinaldo Pis Diez**

En este artículo el Dr. Reinaldo Pis Diez explica los tipos de clusters que existen y cómo openMosix y Condor nos proveen una solución al problema de "load balancing" (balance de carga). En un artículo previo de esta edición (ver página 48) se explica cómo implementar un cluster "Beowulf", usados fundamentalmente en High Performance Computing.

Uno de los objetivos del *clustering* es distribuir eficientemente entre los componentes del cluster, fundamentalmente entre aquellos que se encuentran libres, la carga ocasionada por los requerimientos de los usuarios.

Se acepta en la actualidad que existen tres tipos de clusters [1]:

1. Cluster "a prueba de fallas" es aquel conformado por dos o más máquinas conectadas en red de forma tal que constantemente se controla la actividad de todos los servicios prestados. En cuanto un servicio se detiene en una máquina por cualquier causa, otro componente del cluster intenta reemplazarlo inmediatamente. Esta configuración se conoce también como "redundancia".

2. En un cluster con balance de carga los requerimientos entrantes son enviados al componente más aliviado del mismo. Un cluster de este tipo es también un cluster a prueba de fallas, sólo que con un mecanismo de balance de carga activado.

3. Un cluster para computación de alto rendimiento, conocido como "Beowulf", es diseñado para soportar los importantes requerimientos de cálculo numérico de centros de investigación. También posee mecanismos de balance de carga y soporta entornos de programación especiales que permiten procesar información en paralelo.

En términos del tipo de tareas a desarrollar, un cluster puede ser clasificado según los siguientes criterios:

1. Las tareas que implican el manejo de una gran cantidad de datos durante un prolongado período de tiempo se denominan genéricamente "computación de alta producción" (High-Throughput Computing o HTC) y requieren de un cluster HTC.

2. Las tareas que requieren inmensos recursos computacionales en tiempos relativamente más cortos se definen como "computación de alto rendimiento" (High-Performance Computing o HPC) y requerirán un cluster HPC.

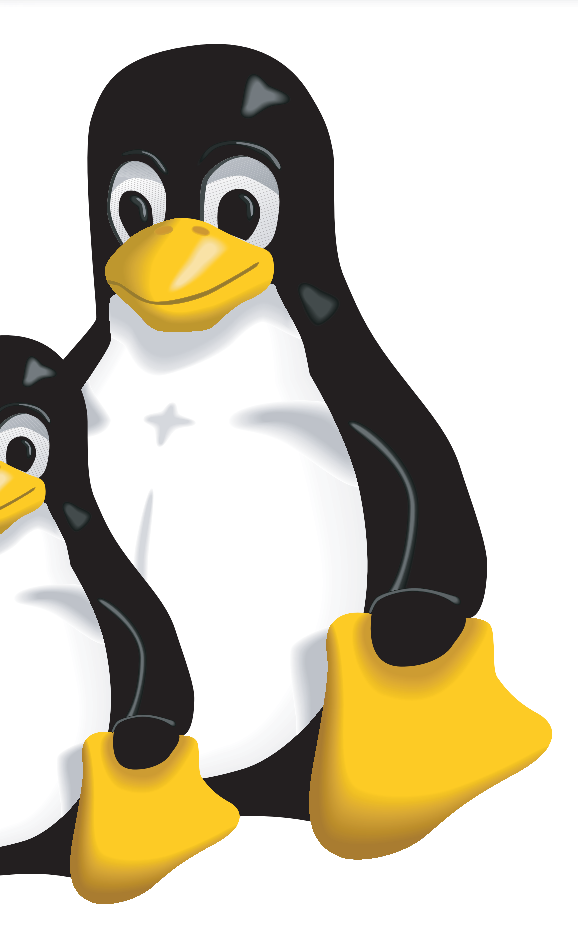
Los inicios de la tecnología Mosix se remontan a los comienzos de la década de 1980. Hacia 1997, y debido a un rumbo netamente comercial tomado por los desarrolladores de Mosix, Moshe Bar decide llevar a cabo el proyecto openMosix bajo licencia GPL2.

OpenMosix es un parche a nivel del kernel de Linux compatible con plataformas de tipo IA32 [2]. Un sofisticado algoritmo basado en premisas de la economía de mercado es utilizado para determinar cuál nodo es el indicado para transferirle un proceso de forma de optimizar el balance de carga total del cluster. openMosix tiene un importante conjunto de herramientas de configuración, incluyendo la interfaz gráfica dada por openMosixview [3] que simplifica notablemente la administración de un cluster.

Una interesante característica de openMosix es la denominada "autodiscovery" que permite detectar nuevos nodos en un cluster y actualizar automáticamente la configuración del mismo para reflejar los cambios producidos.

openMosix

openMosix es un parche a nivel del kernel de Linux compatible con plataformas de tipo IA32. Además tiene un importante conjunto de herramientas de configuración, incluyendo la interfaz gráfica dada por openMosixview y es particularmente útil en sitios donde los usuarios deben llevar a cabo múltiples instancias relativamente rápidas de uno o más programas.



Clusters: un poco de historia

• Décadas del '50 y '60: la Fuerza Aérea norteamericana construye una red de computadoras (funcionando con válvulas) como método de defensa contra un eventual ataque nuclear de la entonces Unión Soviética. Denominaron a esa red SAGE.
• Mediados del '80: la Digital Equipment Corporation acuñó el término "clusters" para bautizar su sistema de minicomputadoras VAX.
• Comienzos de los ´90: gracias a un bajo costo de las PCs, a un rápido desarrollo de la tecnología Ethernet y a la aparición de Linux las PCs pueden ser usadas como "unidades de construcción" de un cluster.
• 1994: el Centro Espacial Goddard de la NASA presentó un cluster de 16 PCs (usando procesador Intel486 bajo Linux). Potencia máxima de cálculo de 70 megaflops (1 megaflops = 1 millón de operaciones de punto flotante por segundo), similar a la brindada por sistemas comerciales cuyo valor era 10 veces superior. Denominaron al cluster "Beowulf", en honor al rey anglosajón del medioevo que derrotó al monstruo Grendel arrancándole un brazo.

A diferencia de un cluster Beowulf, donde lo usual es ejecutar procesos de larga duración, openMosix es particularmente útil en sitios donde los usuarios deben llevar a cabo múltiples instancias relativamente rápidas de uno o más programas. Sin embargo, aplicaciones en paralelo bajo Message Passing Interface, MPI [4], pueden perfectamente llevarse a cabo en un cluster openMosix: si un subproceso es iniciado en el nodo 1 y en determinado momento el nodo 2 presenta menor carga, entonces el subproceso es migrado a este último nodo de una forma que resulta totalmente transparente al usuario.
Condor es un sistema de administración inicialmente diseñado para HTC, donde lo esencial es que el proceso finalice felizmente dado el gran tiempo de cálculo necesario [5].
Al igual que openMosix, no es necesario modificar los programas existentes para poder utilizarlos en un cluster administrado por Condor. Sin embargo, si es posible reconstruir el o los ejecutables aprovechando un conjunto de funciones de biblioteca de Condor, entonces dos importantes ventajas salen a la luz: la generación de checkpoints y llamadas de sistema remotas.

Condor

Condor es un sistema de administración donde no es necesario modificar los programas existentes para poder utilizarlo. Asimismo es extremadamente flexible en su configuración, permitiendo incluso que computadoras de escritorio formen parte de un cluster desde la hora 0 del día A hasta la hora 4 del día B.

Un checkpoint es un archivo que contiene la información completa del estado de un proceso. De esta manera, cuando ocurre un corte de luz o la computadora debe ser reiniciada por alguna razón, el proceso es relanzado desde el último estado registrado en el archivo checkpoint. Por otro lado, cuando el nodo en el que originalmente es iniciado un proceso sale temporariamente de servicio, Condor puede migrar a otro componente del cluster tanto el proceso como el archivo checkpoint asociado para continuar la tarea.
Cuando un proceso es iniciado por un usuario en un componente dado de un cluster, Condor puede migrarlo a otro nodo (nodo remoto) pero mantener los archivos de entrada y salida en el nodo original. De esta forma, es Condor quien se encarga de las llamadas remotas al sistema en lugar del sistema operativo del nodo remoto. Una notable ventaja de este mecanismo es que el usuario ni siquiera necesita una cuenta en el nodo remoto ya que es Condor quien se encarga de todo el proceso.
Condor es extremadamente flexible en su configuración, permitiendo incluso que computadoras de escritorio formen parte de un cluster pero sólo a partir de las 18 horas de cada día y hasta las 8 de la mañana del día siguiente y durante la totalidad de los fines de semana.
Habiendo reseñado las principales características de openMosix y Condor, haremos algunas observaciones que deberían aclarar la verdadera utilidad de estos dos productos.
Básicamente, tanto openMosix como Condor son "administradores de colas". Ninguno de los dos ofrece un entorno de programación en paralelo. Para ello, es necesario utilizar Parallel Virtual Machine, PVM [6], o el ya mencionado MPI, e introducir las modificaciones al software original como llamadas a funciones de biblioteca de aquellos. Ni openMosix ni Condor pueden reemplazar estas utilidades.
Se suele indicar que no es necesario recopilar un programa si se instala openMosix o Condor en un cluster. Esto es cierto para cualquier administrador de colas, no sólo para los productos que se discuten en este artículo. Por otro lado, si un programa es paralelizado, usando por ejemplo las funciones provistas por MPI, entonces el ejecutable deberá reconstruirse independientemente de la naturaleza del administrador de colas instalado en el cluster. Ni openMosix ni Condor pueden evitar la recopilación de la versión paralelizada del programa original.
Se cita también como una ventaja de estos productos que cuando un usuario ejecuta un programa, puede olvidarse del camino que seguirá el mismo hasta ser notificado de su finalización. En realidad, cualquier administrador de colas efectúa esta tarea y no es una característica única de openMosix ni de Condor.
Para finalizar, digamos que tanto openMosix como Condor son dos excelentes productos que pueden funcionar muy bien como una solución al problema del balance de carga en clusters para HTC o HPC. No obstante, no debe perderse de vista que son fundamentalmente administradores de cola inteligentes que no proveen por sí mismos un entorno de programación o de trabajo en paralelo aunque sí son capaces de llevar a cabo procesos que se reparten entre varios nodos. ●

Referencias

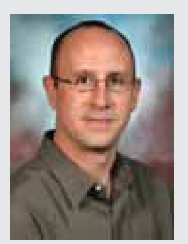
[1] "Introducing openMosix", K. Buytaert. http://www.linuxdevcenter.com/pub/a/linux/2004/02/19/openmosix.html
[2] http://openmosix.sourceforge.net/
[3] http://www.openmosixview.com/
[4] http://www-unix.mcs.anl.gov/mpi/mpich/
[5] http://www.cs.wisc.edu/condor/
[6] http://www.csm.ornl.gov/pvm/

Viridian

Por: **Bill Hilf**
Título original: **Porqué la virtualización es tan popular.**

Cada vez más gente está hablando sobre la virtualización y sus beneficios. Bill Hilf, General Manager of Platform Strategy de Microsoft, explica porqué se da esto y cuenta los avances de Microsoft con respecto al tema.

Quién es Bill Hilf



Bill Hilf es el General Manager of Platform Strategy de Microsoft. Su principal objetivo es lograr la delantera mediante estrategias a largo plazo en el servidor de Windows y en las herramientas de organización. En este aspecto, Hilf generalmente intenta crear y mantener un canal directo entre la comunidad y, de esta forma, generar un feedback.

Antes de unirse a Microsoft, Hilf estuvo al frente de la estrategia técnica de Linux/Open Source Software de IBM para los mercados emergentes.

Asimismo, durante 12 años estuvo involucrado con el desarrollo de un Software Open Source (OSS).

Previo a IBM, fue vicepresidente de ingeniería de eToys, donde ayudó a crear uno de los principales negocios existentes en la web.

Recién volví de la conferencia Linux World (Mundo Linux) en San Francisco y la virtualización fue, nuevamente, el tópico principal. Muchos de los que están afuera de la esfera de la tecnología (donde a nosotros nos gusta hablar con anacronismos raros y pa-labras corporativas) se podrán preguntar porqué Microsoft y muchos otros no pueden parar de hablar sobre la virtualización. Basta con ir a cualquier conferencia sobre IT y es muy probable que haya al menos una charla sobre el tópico de la virtualización. Estas charlas usualmente están acompañadas por diagramas de *marketecture*, con gráficos tipo bloques de Lego con imágenes de diferentes sistemas operativos todos corriendo encima de un solo servidor físico. Habiendo trabajado para IBM estoy bastante familiarizado con la idea de la virtualización, muchas veces llamada, en el lenguaje usado en IBM, 'logical partitioning'. De todas formas, la razón por la cual hoy existen muchas discusiones alrededor del tema de la virtualización es porque está mucho más accesible y a un mejor precio que en el pasado. Intel y AMD han mejorado sus microprocesadores de forma de hacerlos aptos para la virtualización (en el pasado, los administradores de las máquinas virtuales tenían que hacer todo tipo de trucos para lograr la virtualización sobre un intruction-set x86). Esto ha permitido que los programadores de software de máquinas virtuales puedan crear poderosas tecnologías, muchas veces llamadas *hypervisors*, que pueden residir en el mismo sistema operativo permitiendo una virtualización de un sistema operativo sobre un "host" de forma mucho más eficiente y confiable.

¿Es un proyecto científico moderno o existe un uso real de estas cosas? Déjenme darles un simple ejemplo de cómo lo hemos usado en nuestro laboratorio Open Source. Nosotros implementamos diferentes tipos de distribuciones Linux y sus varias versiones, y sobre diferentes arquitecturas de hardware para su testeo y análisis, probablemente más de 50. Típicamente, se podría usar un solo servidor (incluso hasta una sola PC) para cada sistema operativo, lo que significaría aproximadamente 50 máquinas diferentes. Cada una de

Hypervisor

Hypervisor es una tecnología que permite correr, al mismo tiempo, diferentes sistemas operativos (sin modificar) en una misma computadora. Es una extensión de un término anterior "supervisor", que se aplicaba a kernels de sistemas operativos.

Los hypervisors fueron originalmente desarrollados a principios de los años '70 cuando, para reducir costos, se consolidaban varias computadoras aisladas de diferentes departamentos de la empresa en una sola y más grande -el mainframe- capaz de servir a múltiples sectores. Al correr múltiples sistemas operativos a la vez, el hypervisor permite una consolidación dando robustez y estabilidad al sistema; aún si un sistema operativo colapsa, los otros continúan trabajando sin interrupción.

La primer computadora diseñada específicamente para virtualización fue el mainframe IBM S/360 Modelo 67. Esta característica de virtualización ha sido un Standard de la línea que siguió (IBM S/370) y sus sucesoras, incluyendo la zSerie actual.

La necesidad de consolidar los diferentes servidores de hoy y de una administración simplificada han hecho renovar el interés en la tecnología de "hypervisors". La gran mayoría de los vendors Unix, incluyendo Sun Microsystems, HP, IBM y SGI han estado vendiendo hardware virtualizado desde los años 2000. Estos sistemas son eficientes pero extremadamente costosos.

Uno de los primeros hypervisors de PCs fue desarrollado a mediados de los '90 y se llamó VMware. La arquitectura x86 usada en la mayoría de los sistemas de PC es particularmente difícil de virtualizar. Pero las grandes compañías, como Intel, están incorporando extensiones que redireccionan las partes ineficientes o deficientes de virtualización de x86, proporcionando un apoyo adicional al hypervisor. Esto permite un código de simple virtualización y una mejor performance para una virtualización completa.

(Contribución de NEX IT Specialist a wikipedia.org)

Carmine

El futuro de virtualización en Microsoft

esas máquinas requiere energía, nuevas partes, mantenimiento, y demás. Los costos aumentan rápidamente; he visto que en algunos data centers la energía y el mantenimiento pueden significar más de la mitad del costo total de las operaciones llevadas a cabo año tras año. En nuestro laboratorio podemos correr casi todas esas distribuciones Linux en un servidor, un servidor HP Opteron de cuatro vías con 8GB de memoria y mucho disco. Esto es para testeo. Se entiende que no se podrían correr tantas imágenes virtuales en producción y gran cantidad de carga. Resultado final, ahorré tiempo y dinero (particularmente en la gestión de sistemas).

También he hablado con clientes que están usando la virtualización para la recuperación de desastres y escenarios de copias de seguridad. Nuevos escenarios de implementación como en un call center o las sucursales de una oficina pueden ser 'instalados' con imágenes virtualizadas en una fracción del tiempo que toma instalar un servidor tradicional. Y escenarios donde los grupos de testeo y la calidad puedan hacer pruebas del hardware y del software a través de una docena de tipos de configuraciones de sistemas operativos. IDC pronosticó que el 45 por ciento de los nuevos servidores comprados durante este año serán virtualizados.

La virtualización es una parte crítica de la estrategia de Microsoft y hemos estado un tiempo en este negocio con nuestra Virtual PC y con Virtual Server 2005. Hoy en día, Virtual Server 2005 R2 está disponible para ser bajado en forma gratuita. También hemos abierto ciertas especificaciones de nuestro Virtual Hard Disk (VHD) con Virtual Server 2005. Estas especificaciones se pueden usar para aprender cómo acceder (leer y modificar) al almacenamiento de datos en una VPC o a un disco duro virtual en un Virtual Server. El formato VHD está disponible bajo una licencia sin costos de derechos de autor (licencia Royalty-free).

En la actualidad estamos haciendo grandes inversiones con nuestro hypervisor 'Viridian' y con el proyecto del System Center Virtual Machine Manager ('Carmine'). Estos son los nombres para nuestro hypervisor de virtualización Windows Server Lonhorn y para el administrador de maquinas virtuales, respectivamente. La versión beta del System Center Virtual Machine Manager también está disponible para ser bajada. Por lo que he visto hasta aquí en el desarrollo de estos productos, éste será un excelente software. Si prefiere puede chequear el post de Mike Neil en Port2 [1] sobre cómo anunciamos y demostramos mucho de lo aquí expuesto en el WinHec este año – donde también hay un video que muestra la demostración de Bill Gates sobre virtualización realizada en el WinHec. [2]

En relación con esto, recientemente hemos anunciado una alianza entre Microsoft y XenSource. XenSource es la compañía alrededor del proyecto de open source Xen – la principal tecnología de virtua-lización de Linux. Peter Levine, CEO de XenSource comentó esta alianza en su conferencia en el Linux World de San Francisco. Junto con XenSource se trabajará para permitir una óptima virtualización entre Windows y Linux, lo que es de gran importancia para los clientes que viven en un entorno heterogéneo buscando la consolidación de los servidores y buscando ventajas de los nuevos escenarios –como describí anteriormente – en el futuro. Este trabajo formará parte de nuestro plan asociado al servidor Longhorn. Personalmente estoy muy excitado con esta sociedad, la cual es un indicador de cómo pensamos, a largo plazo, el mapa de la interoperabilidad.

Están sucediendo muchas cosas en el área de la virtualización y pienso que es uno de los cambios más importantes de nuestra industria. Seguramente se exagerará y habrá todo tipo de promociones al respecto, típico del momento en el que estamos. Sin embargo personalmente he ahorrado tiempo y dinero en mi propio laboratorio y he hablado con clientes que han encontrados ventajas similares. Efectivamente, tiempos excitantes. ●

Traducción realizada por NEX IT Specialist. Extraido de la Web-page de Port 25. http://port25.technet.com/

Xen, XenSource y paravirtualización

Xen es un software Open Source, un hypervisor para x86, Intel, Itanium y PowerPC, que posibilita la ejecución de múltiples sistemas operativos invitados en el hardware de una misma computadora. Esta forma de virtualización es obtenida mediante una técnica llamada paravirtualización. Xen provee un aislamiento seguro, control de los recursos, calidad garantizada de los servicios y la posibilidad de migración de las máquinas virtuales; y su arquitectura de poco peso hace posible que todos los sistemas operativos x86 soporten la paravirtualización. Xen se desarrolló en el Laboratorio de Computación de la Universidad de Cambridge (www.cl.cam.ac.uk/).

XenSource es una empresa que ofrece una plataforma de virtualización, a través de la cual se desarrollan, comercializan y se venden soluciones adicionales y servicios basados en el hypervisor de Xen, incluyendo soluciones de infraestructura de virtualización.

Para más información no dude en visitar su web http://www.xensource.com/

Lectura Adicional

- **www.microsoft.com/windowsserversystem/virtualserver**
- **www.microsoft.com/whdc/system/platform/virtual/default.mspx**
- **www.microsoft.com/systemcenter/scvmm/default.mspx**

Referencias

[1] http://blogs.technet.com/virtualization/archive/2006/05/23/429897.aspx
[2] http://www.microsoft.com/events/executives/billgates.mspx

Business Intelligence y Data Mining

NOTA 6

#1 El negocio del Software: Caso de Éxito Snoop Consulting
#2 Gente, Procesos y Métodos (Métodos ágiles y disciplinados)
#3 Modelado y diseño de sofware (De la nada al software, hoy)
#4 Control de Calidad de Software
#5 JAVA vs. .NET
#6 Data Warehouse, Bussiness Intelligence y Data Mining
#7 Coordinando procesos de negocios con BPEL

Por: **Gustavo Walter** y **Fernando Das Neves**
Snoop Consulting

Desde principios de los 90 las aplicaciones de BI y Data Mining han evolucionado dramáticamente en muchas direcciones. El motivo de este artículo es dar un panorama general y actualizado de las componentes que forman parte de una solución de *Business Intelligence y Data Mining* y qué se puede lograr con tal solución.

Business Intelligence

Últimamente se ha hecho uso y abuso de estas palabras. Si buscamos una definición de Business Intelligence (BI) podríamos decir que es un conjunto de estrategias y herramientas enfocadas a la administración y creación de conocimiento mediante el análisis de datos existentes en una organización o empresa.

El exceso de información no es poder, pero el conocimiento sí. Con demasiada frecuencia la transformación y el análisis de la información y los datos que las compañías generan se convierte en un verdadero problema y, por lo tanto, la toma de decisiones se vuelve desesperadamente lenta.

Las tecnologías de BI intentan ayudar a las personas que toman las decisiones a entender los datos más rápidamente para que puedan tomar mejores y más rápidas decisiones y, así, mejorar sus acciones a futuro.

¿Qué es BI?

Para entender un poco más acerca de BI podríamos hacer una analogía con una "Refinería de datos". Hay que pensar en una refinería petrolera que esta diseñada para tomar petróleo crudo, procesarlo y generar una gran variedad de productos. De la misma manera, una solución de BI toma datos, los procesa y genera una gran variedad de productos de información.

La figura 1 ejemplifica una "Refinería de Datos". Veamos los productos de información que estaría generando:



- **INFORMACIÓN:** Una herramienta de ETL (Extract, Transform, Loading) extrae datos de los diferentes Sistemas Transaccionales (OLTP) de una organización, los integra y almacena en una base de datos dedicada para estos fines (Datawarehouse). Por ejemplo, un data warehouse puede integrar información del cliente de cinco diferentes sistemas transaccionales (Ordenes de Compra, Ordenes de Servicio, Ventas, Entregas y Programas de Lealtad del Cliente). Este proceso de extracción e integración transforma los datos en un nuevo producto: INFORMACIÓN.

- **CONOCIMIENTO:** Los usuarios provistos de herramientas de análisis como OLAP o Herramientas de Minería de Datos analizan la información almacenada en el data warehouse. En su análisis identifican tendencias, patrones y excepciones, lo cual les permite transformar la información en CONOCIMIENTO.

- **PLANES:** Con el conocimiento obtenido, los usuarios lo transforman en REGLAS. Éstas pueden ser muy simples (por ejemplo, "Ordenar 50 unidades del material X cada vez que el inventario llegue a 25 unidades") o reglas muy complejas generadas por algoritmos estadísticos o modelos matemáticos. Luego, los usuarios crean PLANES que implementan estas reglas.

- **ACCIONES:** Los planes generan ACCIONES que serán ejecutadas en la organización, por ejemplo, un plan que defina qué productos ofrecer a un cliente de acuerdo a la predicción de aceptación que tendrán.

- **EVENTOS:** Estas acciones generan EVENTOS que afectan a la organización, son asimilados y transformados en transacciones que son ejecutadas sobre los sistemas transaccionales y convertidas en datos que entrarán nuevamente a la "refinería de datos" permitiéndonos evaluar la efectividad de nuestras ACCIONES.

Esto nos permite entrar en un círculo virtuoso que crea una organización de APRENDIZAJE que distingue a las compañías "inteligentes" y que utilizan eficientemente su proceso de BI.



Fig. 1 Esquema de los componentes conceptuales de una solución de BI.

Componentes de una Solución de Business Intelligence

En una solución de BI participan muchos componentes (Figura 2):

¿Qué es un DataWarehouse?

El término Datawarehouse fue acuñado por Will Inmon quien lo definió como "una colección de datos orientados al tema, integrados, no volátiles e historiados, organizados para el apoyo de un proceso de ayuda a la decisión".

- **Integrado:** Los datos almacenados en el Datawarehouse deben integrarse en una estructura consistente, las inconsistencias existentes entre los diversos sistemas operacionales deben ser eliminadas.

- **Temático:** Los datos se organizan por temas para facilitar su acceso y entendimiento por parte de los usuarios finales. Por ejemplo, todos los datos sobre clientes pueden ser consolidados en una única tabla del Datawarehouse. De esta forma, las peticiones de información sobre clientes serán más fáciles de responder dado que toda la información reside en el mismo lugar.

- **Histórico:** El tiempo es parte implícita de la información contenida en un Datawarehouse. En los sistemas operacionales, los datos siempre reflejan el estado de la actividad del negocio en el momento presente. La información almacenada en el Datawarehouse sirve, entre otras cosas, para realizar análisis de tendencias. El Datawarehouse se carga con los distintos valores que toma una variable en el tiempo para permitir comparaciones.

- **No Volátil:** El almacén de información de un Datawarehouse existe para ser leído y no modificado. La información es por tanto permanente, lo que significa la actualización del Datawarehouse y la incorporación de los últimos valores que tomaron las distintas variables contenidas en él sin ningún tipo de acción sobre lo que ya existía.

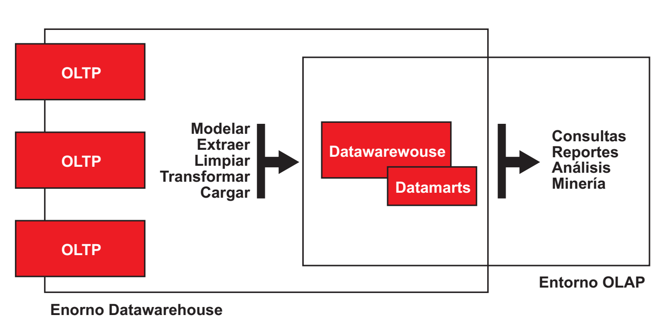


Fig. 2 Componentes de una solución de BI

Obviamente la definición anterior debe tomarse como la definición "pura" sobre datawarehouse. Sin embargo luego de diez años algunos términos han sido manejados según las necesidades y capacidades del mercado, dando origen a nuevos conceptos como **Datamart** para referirse a un área especifica dentro del warehouse corporativo.

¿Cómo se modela un DW?

El modelo que soporta la información que contiene un DW se encuentra diseñado, estructurado e implementado con la finalidad del análisis y navegación de los datos. Se entiende por navegación o **drilling** de los datos la posibilidad de ver información correspondiente a diferentes contextos o entornos, por ejemplo, analizar las ventas anuales y poder "abrirlas" por sucursal, después analizar en más en detalle para ver cómo se discriminan las ventas por cada producto, etc.

¿Cómo se diseña un DW?

Diseño "Estrella" (Figura 3): Es la técnica de diseño más popular usada para un datawarehouse. Es un paradigma en el cual un único objeto en el centro (conocido como tabla de hechos) está conectado con otros objetos circundantes llamados tabla de dimensiones formando una estrella.

Diseño "copo de nieve" (Figura 4): es una extensión del esquema estrella donde cada punta de la estrella se explota en más puntas y su denominación se debe a que el diagrama del esquema se asemeja a un copo de nieve. Los esquemas de este tipo tienden a normalizar dimensiones para eliminar redundancia. Los datos de las dimensiones se agrupan en múltiples tablas en lugar de una tabla grande.

Ambiente OLAP

En el lado derecho de la figura 1 ubicamos al entorno OLAP (el entorno analítico) que es el dominio de los usuarios que explotaran la información provista por una solución de BI. Las características que debe tener una herramienta de BI se podría resumir en:

• **Accesibilidad a la información.** Los datos son la fuente principal de este concepto. Lo primero que deben garantizar este tipo de herramientas y técnicas será el acceso de los usuarios a los datos con independencia de la procedencia de estos.

• **Apoyo en la toma de decisiones.** Se busca

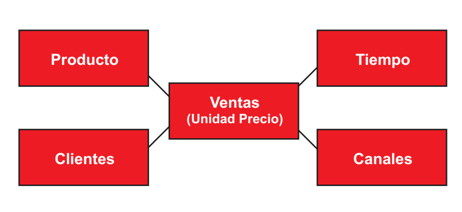


Fig. 3 Modelo Estrella

ROLAP vs MOLAP

Según la implementación seleccionada, los datos son almacenados en forma relacional, comúnmente llamado ROLAP, los cuales son implementados en bases de datos RDBMS respetando ciertos estándares a nivel de definición y consulta de datos o en formato multidimensional llamados MOLAP. Las bases de datos multidimensionales son arquitecturas propietarias definidas por cada proveedor que son frecuentemente actualizadas desde bases de datos relacionales.

ir más allá en la presentación de la información, de manera que los usuarios tengan acceso a herramientas de análisis que les permitan seleccionar y manipular sólo aquellos datos que les interesen.

• **Orientación al usuario final.** Se busca independencia entre los conocimientos técnicos de los usuarios y su capacidad para utilizar estas herramientas. Armando catálogos acorde al vocabulario que maneja cada sector de los tomadores de decisión.

De acuerdo al nivel de complejidad las soluciones de BI se pueden clasificar en:

• Consultas e informes simples (Queries y reports)
• Aplicaciones OLAP (On Line Analytic Processing)
• Data Mining o minería de datos

Desde hace un tiempo han venido surgiendo nuevas tecnologías que permiten explotar la información que reside en una solución de BI, por ejemplo Data Mining, Scorecard, cuadros de mando, planning, etc, los cuales han hecho más complejo el BI.

¿Por qué fallan las soluciones de BI?

Los proyectos de BI fallan debido a que:

• Las empresas fallan en reconocer que los proyectos de BI son iniciativas de negocios interorganizacionales y en entender que dichas iniciativas difieren de las típicas soluciones independientes.
• Existe una falta de compromiso por parte de los *sponsors* (los cuales tienen autoridad en la empresa).
• Se tiene poca disponibilidad de los representantes de negocios.
• Hay ausencia de un personal disponible y habilidoso.
• Existe un mal concepto del software de BI.
• No se trabaja bajo una estructura detallada.
• No existe un análisis del negocio o estandarización.
• No existe una apreciación del impacto que causan los datos de mala calidad en la rentabilidad del negocio.
• No se entiende la necesidad del uso de un metadatos para que todos hablen el mismo idioma.
• Métodos y herramientas no alineadas.

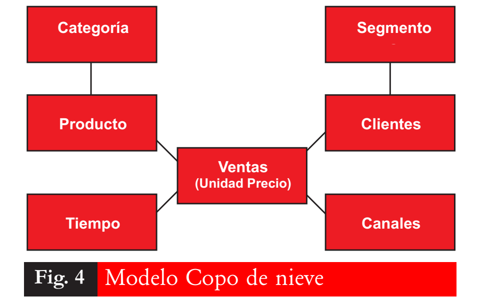


Fig. 4 Modelo Copo de nieve

Riesgos de BI

El uso apropiado de las herramientas de BI puede marcar la diferencia entre la vida y la muerte de muchas empresas, entre el estancamiento y el crecimiento, entre el servicio de mala calidad y el servicio al cliente personalizado, y entre la relación óptima con los proveedores y la pérdida de los beneficios de trabajar con ellos y con otros socios de negocios. Por todo ello una buena solución de BI es fundamental.

El mayor riesgo tecnológico es que la tecnología esta cambiando rápidamente. Naturalmente, las nuevas tecnologías tienen algo de riesgo hasta que son probadas completamente. Algunos de los grandes riesgos relacionados con el uso de las herramientas de BI están basados en los datos que son usados pero no transformados apropiadamente. Debido a que en el ámbito de los negocios las empresas muy frecuentemente escogen sus propias herramientas de BI, una empresa puede terminar con múltiples herramientas, o múltiples *data marts* con datos que no están claramente definidos o con meta datos que no son compatibles entre sí. Esto puede inducir a tener diferentes conclusiones acerca de los mismos datos.

Data Mining

La gente no actúa al azar. La lista del supermercado de la gente se parece mes a mes; hay una relación entre el aumento de actividad económica y el aumento de la demanda de ciertos tipos de crédito; la mayoría de los amigos de una persona comparten intereses políticos con la persona.

Casi cualquier acción de una persona hoy en día deja un registro electrónico: cada ítem comprado en el supermercado, cada pedido de crédito, cada email mandado a un amigo. Sin embargo la cantidad de información generada es tal que es imposible de ser comprendida por una persona. Lo que necesitamos es una manera de describir automáticamente la relación entre datos, para que aún si no los describen totalmente, capturen lo suficiente de las relaciones como para poder comprenderlos mejor, explicarlos y, de ser posible, predecir cómo van a cambiar esos datos en el futuro. Las técnicas que toman datos y encuentran patrones y tendencias en los datos, a partir de tendencias regulares, se engloban dentro del área de informática conocida como *machine learning*. A la aplicación de machine learning a grandes bases de datos se la llama data mining, como una analogía entre buscar una pequeña cantidad de minerales valiosos entre toneladas de piedras, y la de extraer una serie de patrones significativos a partir de millones y millones de registros en bases de datos.

Las técnicas de data mining pueden usarse para describir y clasificar los datos que están siendo analizados (análisis descriptivo) o para predecir o clasificar datos futuros (análisis predictivo). La mayoría de los proyectos de data mining involucran una combinación de ambos tipos de análisis: entender qué está pasando, y hasta donde se pueden sacar conclusiones a partir de los datos, para luego poder hacer pronósticos de tendencias o clasificar nuevos datos según llegan ¿Cómo hace uno para testear si un modelo es válido a futuro, si el futuro es desconocido (simplemente por ser futuro)? La manera es dividir el conjunto de datos disponibles en dos subconjuntos: el de entrenamiento y el de prueba. Los parámetros de los métodos de data mining se calculan (*se entrenan*) a partir de los datos del conjunto de prueba, y la validez de esos parámetros se comprueba aplicando el modelo definido con sus parámetros ya calculados a los datos de prueba, de los que uno sabe las características, y viendo si lo predicho por el modelo coincide con lo esperado. Si eso ocurre, entonces es esperable que también sean capaces de aproximar el futuro cercano. La diferencia entre lo predicho por el modelo y el valor real es el error promedio de validación del modelo, que se puede medir de varias maneras como el error medio cuadrado (MSE), o el error medio absoluto (AME).

¿Cómo predecir? ¿Qué es posible predecir? ¿Con qué supuestos?

Cuando existe una pregunta que puede contestarse con data mining, la calidad de la respuesta depende fundamentalmente de la calidad de los datos; normalmente se tiene menor volumen de datos de los necesarios para conseguir una clasificación o predicción de la calidad deseada. La complejidad de obtener respuestas con data mining se deben a los siguientes factores:

- **¿Cuántas dimensiones (columnas en la tabla) tiene el espacio de datos? Cada campo, ¿tiene valores discretos o continuos?** Al aumentar el número de clases crece expo-

nencialmente el número de ejemplos necesarios para poder obtener una clasificación de calidad. Además, al aumentar el número de dimensiones nuestra intuición de cómo se distribuyen los puntos en el espacio no es válida. Este problema del crecimiento rápido en el número de ejemplos necesarios y el comportamiento no intuitivo de espacios de altas dimensiones es tan importante que se lo llama "La Maldición de la Dimensionalidad".

- **¿Tenemos ejemplos de la clasificación deseada?** Supongamos que queremos aprender una manera de clasificar valores ¿Tenemos ejemplos de valores y de la clasificación deseada para cada uno, o queremos que el sistema identifique (diferencie) grupos de valores relacionados? Si tenemos valores (filas de una tabla en una base de datos) y su clasificación deseada, y queremos que el sistema aprenda cómo asignar una categoría a un valor nuevo, tenemos entonces un problema de *aprendizaje supervisado*, llamado así porque los ejemplos sirven para guiar al sistema sobre lo que debe aprender. Si en cambio sólo tenemos una serie de valores, pero no tenemos ejemplos donde los valores estén asignados a categorías, vamos a necesitar un enfoque de *aprendizaje no supervisado*.
- **¿Cuál es la respuesta buscada?** La respuesta más fácil de conseguir es un sí/no, como por ejemplo "dado su historial, ¿es éste un cliente en riesgo de irse a otro proveedor?". Este tipo de respuestas es más simple de encontrar porque solo hay dos clases ("si" y "no") que aprender a separar. Siguiendo con el ejemplo de los clientes, las dos categorías serían aprender cómo identificar clientes entre fieles y riesgosos a partir de la información de clientes que se han quedado en la compañía luego de quejarse, y de otros clientes que se han ido. Cuando el problema es aprender cómo clasificar ítems en una de múltiples categorías (llamadas en la jerga *clases*), sería muy útil saber cual es la "forma" de las clases: supongamos que dibujamos cada ítem como un punto en una hoja, con un color diferente para cada punto, dependiendo a qué clase pertenece. Si dibujara una curva alrededor de los puntos de cada categoría, ¿los puntos se distribuyen en grupos separables o casi separables aceptando unos pocos puntos que caen del lado equivocado de las curvas? ¿Cuál es la forma más simple en que se los puede separar? Tal vez líneas rectas sean suficientes, tal vez la separación es mucho más compleja y requiere curvas cerradas. En la práctica, la cantidad de dimensiones del espacio de datos es siempre mayor que tres, por lo que no tenemos manera de dibujarlas en tres dimensiones e intuir cómo separarlas. Es por eso que los métodos de separación son una combinación de intuición sobre el espacio de datos, y métodos robustos que puedan aproximar aún los casos complejos.

¿Predecir usando qué?

Lo que obnubila a quien recién se inicia con data mining es la cantidad de algoritmos disponibles. Aquí enumeramos algunos algoritmos y familias de algoritmos, explicando el concepto detrás de cada uno para dar una idea para que la próxima vez que vean los nombres, le suenen familiares.

Árboles de decisión: Los árboles de decisión (como casi todos los algoritmos de clasificación) sirven tanto para clasificar como para predecir. Un árbol de decisión es un árbol donde cada rama representa un criterio de decisión (una pregunta que se contesta con sí o no). Para clasificar un punto basta con bajar por el árbol desde la raíz, siguiendo las preguntas cuya respuesta es "sí" para ese punto, hasta alcanzar un nodo terminal del árbol. Los criterios son decididos por el algoritmo, buscando el atributo (columna) que mejor distinga a las clases. Esto se calcula con algún criterio de optimización (como es maximizando el information gain en el algoritmo de construcción de árboles C4.5 o minimizando el error cuadrático en CARTS). La mayor ventaja de los árboles de decisión es que son fáciles de comprender; una persona puede ver el árbol y entender qué hace y cómo se usa para tomar una decisión. La desventaja es que sólo funcionan con atributos discretos, por lo que si un atributo tiene valores continuos necesita ser discretizado primero en rangos de valores. Algunos algoritmos de árboles de decisión se encargan de esto como un paso previo.

SVM (Support Vector Machines) y SVRM (Support Vector Regression Machines): Estos algoritmos toman una serie de ejemplos clasificados en dos categorías (sí/no, legal/fraudulento, etc.) y busca una curva (decidida por adelantado) cuya separación con los puntos en cada categoría sea no menor a un cierto valor (llamado el *margen*, porque es como un área alrededor de la curva donde no debería haber puntos). Cómo encontrar una separación perfecta es generalmente imposible, SVM permite que el ajuste de la curva incluya algunos puntos dentro del margen, pero que en conjunto estos puntos solo "se meten" en el margen hasta un cierto límite (llamado el *margen blando*). SVM es hoy en día unos de los mejores algoritmos de clasificación conocidos, con performance similar a redes neuronales pero sin la necesidad de "brujería" de parámetros y topología de la red que han sido el problema tradicional de las redes neuronales. SVMR es una variación de SVM usada para predicción (acomodar una curva para que encaje una serie de puntos lo mejor posible, de manera de descubrir cuál es la función que relaciona esos puntos).

Clustering: Éste es el nombre de toda una



familia de algoritmos cuya función es la de encontrar la manera de agrupar ítems, con el objetivo de que ítems en el mismo grupo compartan una o más características. Los algoritmos de clustering no necesitan ejemplos para aprender a separar sino que usan un criterio prefijado para decidir si dos ítems son suficientemente similares o no. La ventaja de los algoritmos de clustering es que, sin que el usuario tenga que saber nada acerca de los datos el algoritmo puede ayudar a entender los datos de alto nivel, agrupándolos en unas pocas categorías. La desventaja de clustering es que en realidad los clusters no existen; siempre hay muchas maneras de que una persona pueda agrupar datos en categorías. La agrupación tiene sentido o no dependiendo de que lo que el usuario busca coincida con el criterio de similaridad y agrupación usado por el algoritmo.

Reglas de Decisión: Son algoritmos sin aprendizaje que permiten, a partir de una tabla, encontrar reglas que relacionan valores. El ejemplo clásico es el de analizar una tabla donde cada columna es un ítem en el super-



mercado y cada fila la lista de ítems que había en el carrito de supermercado de un cliente. Aplicando reglas de decisión a esta tabla es posible encontrar relaciones del tipo "si es hombre y compra esto y esto, entonces es probable que compre aquello y aquello otro también", que sirven para que el supermercado coloque a todos esos artículos cercanos, de manera de que el cliente los encuentre fácilmente y compre por impulso.

Conclusiones

Las organizaciones deben usar BI para sostener las inversiones realizadas en años previos en aplicaciones empresariales que han derivado en el uso de enormes cantidades de datos. De esta manera BI valida, mide y maneja nuevas oportunidades e inversiones en nuevos negocios.

Las técnicas de *Business Intelligence* y Data Mining posicionan a una compañía para generar el mayor valor de las líneas de negocios existentes y anticipar nuevas oportunidades. Asimismo, las soluciones de *BI* y Data Mining pueden ayudar a la compañía a reducir los ciclos de desarrollo de productos, agilizar operaciones, afinar campañas de marketing y mejorar relaciones con clientes y proveedores, todo lo cual significan menores costos y mayores márgenes de utilidad. ●

Microsoft .net

.NET 2.0 Framework Características avanzadas

El .NET Framework 2.0, además de mejorar el rendimiento con respecto a las versiones anteriores, agrega nuevas funcionalidades como tipos genéricos, clases parciales, tipos anulables, alias externos, método anónimos, etc.



Gabriela Marina Giles

- **Microsoft .NET Senior Trainer** CentralTech GOLD CERTIFIED Partners
- **Presidenta de Desarroll@doras** Grupo de usuarios de Tecnologías .NET www.desarrolladoras.org.ar

Nota #2 de 6 Serie

1- .NET Framework

2- Características avanzadas del .Net Framework 2.0

3- ASP.NET ¿Qué hay de nuevo para el diseño de Aplicaciones Web?

4- ASP.NET Seguridad y Manejo de estado

5- ADO.NET Introducción – Creando un proyecto de datos

6- Web Services nhancements

Genéricos

Generics permite a los desarrolladores probar código, escribirlo una vez y reutilizarlo para una variedad de diferentes tipos de datos. Las funciones de lenguaje conocidas como genéricas actúan como plantillas que permiten declarar y definir interfaces, estructuras, clases, métodos y delegados con parámetros de tipo genérico, en lugar de tipos específicos. Cuando se utiliza el genérico quedan especificados los tipos reales. GENERICS tiende hacer más eficiente que las colecciones de Objetos.

Características principales de Generics:
1. Aumento de performance, NO NECESITA HACER BOXING Y UNBOXING O DOWNCASTS.
2. Instanciados en run-time, no en compile-time.
3. Verificados en la declaración, no en la instanciación.
4. Información exacta del tipo en run-time.
5. Funcionan para tipos por referencia y por valor.
6. No es necesario hacer el Cast del ítem solicitado de una colección.
7. Reduce la necesidad de conversiones explícitas.
8. Resulta en un tipo de código seguro, limpio y rápido.
9. Habilita la creación de modelos de códigos.
10. Todos los lenguajes .NET pueden producir y usar Generics.

Veamos el siguiente código (C#) de ejemplo, cuando generábamos una cola que admitiese objetos de cualquier tipo en versiones anteriores del .Net Framework:

```
//Definición de una clase que trabaja con
elementos y método genéricos definidos como
"Object"
public class Cola
{
    object[] elementos;
    public int NúmeroElementos;
    public void Encolar(object valor);
{…}
    public object Desencolar()
{…}
}
    //Instancia de la clase Cola

    Cola miCola = new Cola();
    miCola.Encolar("Prueba");
    string valorDesencolado = (string) miCola.Desencolar();
```

A la hora de extraer valores de la cola habrá que convertirlos a su tipo real si se quieren aprovechar sus miembros específicos, si por error se introduce un valor distinto a un string (por ejemplo, un int) saltará una InvalidCastException.
Otro problema es su bajo rendimiento, cada vez que se almacene un objeto de un tipo referencia en la cola habrá que convertir su referencia a una referencia a object y al extraerlo habrá que volverla a transformar en una referencia a string, lo que conocemos como boxing y unboxing.

A continuación veremos en mismo ejemplo pero escrito en C# con Genéricos:

```
    public class Cola<T>
{
    T[] elementos;
    public int NúmerodeElementos;
    public void Encolar(T valor)
{…}
    public T Desencolar()
{…}
}
```

La cola no requerirá de ningún tipo de conversión para la extracción de objetos, tan claro como podemos obser-

Cuando se conocen las **fuentes**
no hacen falta preguntas
para hacer la elección **correcta**

.cafeconf 2006

.reservá tu lugar en cafeconf ingresando en http://www.cafeconf.org/

.del 10 al 11 de noviembre encontrate con las fuentes
en Lima 717, Ciudad de Buenos Aires, República Argentina

V Conferencia Abierta de GNU/Linux y Software Libre

.aulaslibres
I Congreso de responsables de aulas informáticas con Software Libre

.sponsors





.organizan









var en el siguiente código:

```
Cola<string> miObj = new Cola<string>();
miCola.Encolar("Prueba");
string valorDesencolado = miObj.Desencolar();
```

Con la implementación de Genéricos, es posible construir una clase de Lista, que trabaje con clases de forma independiente a su tipo, de forma similar al funcionamiento de la característica de Templates del lenguaje C++.

Clases Parciales

En versiones anteriores, el código entero de una clase debía estar contenido en un único archivo. En la nueva versión se permiten clases parciales: fragmentos de una clase contenidos en distintos archivos, es el compilador el encargado de unirlas.
Un nuevo modificador utilizado es "**partial**" cuando queremos definir un tipo en múltiples partes. **partial** debe de aparecer antes de las palabras clave **class, struct o interface**.

En el siguiente código C#, muestra en dos archivos separados la clase "Cliente":

```
public partial class Cliente
{
  private int id;
  private string Nombre;
  private string Domicilio;
  private List<Orders> Ordenes;
}

public partial class Cliente
{
  public void xxxMetodoOrdenes(Order order) {
    orders.Add(order);
  }

  public bool xxxFunctionOrders() {
    return orders.Count > 0;
  }
}
```

Sus aplicaciones son varias. Por ejemplo, es muy útil para el manejo de versiones: un programador puede trabajar en una parte de la clase mientras otro trabaja en otra. También cuando trabajamos con los generadores de código se pueden combinar fragmentos de una clase escritos a mano con los generados automáticamente, por ejemplo una clases parcial DataSet, manteniendo en distintos archivos el código autogenerado de nuestro código.

Tipos anulables
Nullable(Of T) (en C# Nullable<T>).

Los tipos Nullable Types son tipos de datos que nos permiten trabajar con un valor nulo. Por ejemplo, un valor Integer en VB o int en C# no tenían forma de tener un valor nulo. Esto es un tema bastante importante al momento de tratar con base de datos, dado que en estas existen campos con datos nulos.

De la siguiente forma podemos definir un tipo nullable en VB.NET o C#:

```
'VB.NET
Dim nulo as Nullable(Of Integer)
If nulo.HasValue Then
'Tiene un valor
End If

//C#
Int? Nulo;
If (nulo.HasValue)
'Tiene un valor
```

La versión anulable de un tipo valor se representa igual que la normal pero con el sufijo *?*, su valor por defecto será null. Por ejemplo:

```
int? y = 1;
y = null;
```

El uso de ? no es más que una sintaxis abreviada con la que instanciar un objeto del nuevo tipo genérico Nullable<T>. Del ejemplo anterior veremos a continuación el código equivalente, donde este tipo tiene un constructor que admite un parámetro del tipo genérico T:

```
Nullable<int> y= new Nullable<int>(1);
// También valdría
Nullable<int> y = new int?(1);
y = null;
```

El tipo Nullable proporciona dos propiedades a todos los tipos anulables: bool HasValue para indicar si almacena null, y T Value, para obtener su valor. Si una variable anulable valiese null, leer su propiedad Value haría saltar una InvalidOperationException.

Alias externos
Un alias es un conjunto de namespaces. Los alias externos sirven para usar en un mismo proyecto dos assemblies que contienen un nombre completamente calificado incluyendo namespace en común. Supongamos que la clase Program necesita hacer referencia a dos assemblies (assembly1.dll y assembly1.dll). Los dos tienen código como el siguiente ejemplo, donde solo varía el cuerpo del método:

```
namespace Test
{
  public class TestMyClase
  {
    public static void Z() {...}
  }
}
```

El programa puede usar referencias externas:

```
extern alias assembly1;
extern alias assembly2;

class Program
{
  static void Main() {
    assembly1.Test. TestMyClase.Z();
    assembly2.Test. TestMyClase.Z();
  }
}
```

Como no se guardan relación con el nombre de los assemblies, los alias podrían ser cualesquiera, los nombres físicos quedan desacoplados de los lógicos. En la compilación habrá que usar el nuevo switch /r.

```
C:\>csc /r: assembly1= assembly1.dll /r:
assembly1= assembly1.dll test.cs
```

Métodos anónimos

Permiten definir un bloque de Código a ejecutarse por un delegado. Sin tener que crear un método para dicho bloque.

```
miObjeto.miEvento += delegate {
Console.WriteLine ("Evento producido en
miObjeto"); };
```

Delegate

Un delegado proporciona un mecanismo mediante el cual unos objetos pueden solicitar a otros que se les notifique cuando ocurran ciertos sucesos.

Es un tipo de clase cuyos objetos pueden almacenar referencias a uno o más métodos de tal manera que a través del objeto sea posible solicitar la ejecución en cadena de todos ellos.

Por ejemplo, ellos pueden ser una Asociación de código a la carga y descarga de ensamblados, a cambios en el sistema de archivos, a la finalización de operaciones asíncronas, a cambios en bases de datos, etc. ●

Recursos

Para ampliar los conceptos que hemos visto en éste artículo:

MSDN2 Library (Español), System.Collections.Generic:
http://msdn2.microsoft.com/es-es/library/system.collections.generic.aspx

http://www.microsoft.com/spanish/msdn/comunidad/mtj.net/voices/MTJ_1903.asp
http://www.microsoft.com/spanish/msdn/articulos/archivo/030506/voices/C20_default.mspx
http://www.microsoft.com/spanish/msdn/centro_recursos/csharp/default.mspx

.NET Framework y lenguajes punto NET:
http://www.elguille.info/NET/default.aspx

Grupo de usuarios de tecnologías .NET:
www.desarrolladoras.org.ar

26 y 27 de Octubre
SHERATON LIBERTADOR - HOTEL

JORNADAS NACIONALES sobre Tecnologia & Negocios

AGSI 2006

El evento de Tecnología y Negocios del año, que convocará a CIOs e importantes referentes de la disciplina, todos aportando sus experiencias en los temas claves para la gestión de las IT.

Conferencias plenarias dirigidas a CIOs, Gerentes de Sistemas, segundas líneas del área de TI, líderes de proyectos, analistas funcionales, ejecutivos de empresas de TI, consultores de TI, consultores y Ejecutivos de negocios con foco en las TI.

ACREDITESE EN:

WWW.AGSI.ORG.AR

Y QUE PIENSA DE PODER ESCUCHAR LAS REFLEXIONES Y EXPERIENCIAS DE LOS PRINCIPALES REFERENTES DE LA DISCIPLINA:

RAÚL SAROKA
Outsourcing: del exitoso al fracaso, sólo un paso.
HUGO SCOLNIK
Firma digital: como implementarla para lograr procesos de negocio eficientes y seguros
JULIO ARDITA
Experiencias de principales fraudes informáticos.
ANTONIO CASTRO LECHTALER
Presente y futuro de la conectividad
JORGE CROM
Como mirar el negocio desde la conectividad
EDGARDO JURI (representante de Gartner Group)
TI en el Mundo y en Latinoamérica. Tendencias tecnológicas
ALEJANDRO OLIVEROS (Director de investigaciones IDC)
Tendencias tecnológicas del mercado argentino de TI para los próximos años

SILVIO SZOSTAK CIO Cervecería y Malteria Quilmes
Manejo de factores humanos en la incorporación de la tecnología en forma exitosa
FRANCISCO RASTELLI - CIO Bayer
Caso Bayer. Mejores Prácticas en el área de TI (ITIL)
GERARDO DALL'ORSO - CIO Molinos Río de la Plata
Caso Molinos: éxitos y dificultades en la administración del ERP
ALBERTO CHEHEBAR - CIO Aerolíneas Argentinas
Caso Aerolíneas: éxitos y dificultades en la implantación de una solución de Business Intelligence

O ACASO NO ES UNA EXCELENTE INVERSIÓN ESCUCHAR LOS CASOS DE LOS SIGUIENTES CIOs COMPARTIENDO SUS VIVENCIAS, DIFICULTADES Y APRENDIZAJES ?

PATROCINAN: PRICEWATERHOUSECOOPERS



AUSPICIAN Y CONVOCAN:











Consejo Profesional de Ciencias Economicas de la Ciudad Autonoma de Buenos Aires



ORGANIZA:



REALIZA:



INFORMES:

+ JORNADAS06@AGSI.ORG.AR
+ TEL: 54.11.4706.3000

Futuro laboral
como desarrollador

Juan José Evangelista
Coordinador de la oficina
de Desarrollo Profesional
Snoop Consulting

Para aquellos que han terminado de estudiar o los que intentan hacerse camino en la carrera de desarrollador hay variedad de opciones para elegir. En esta nota intento dar un lineamiento de alguna de ellas, teniendo en cuenta las opciones laborales pero también los aspectos técnicos que conlleva cada elección.

Recibirse. Años de un ritmo de estudio, compañeros, prácticos, horarios dispersos, poco sueño, expectativas, proyectos. Se terminó ¿Y ahora qué? Hay que trabajar. Salir a buscar trabajo ¿A buscar trabajo? Sí, es cierto que la única manera de encontrarlo es buscándolo, pero yo preferiría mirarlo de otro modo y llamarlo **salir a ofrecerse**. Buscar pareciera limitarse a recorrer una lista de posibilidades y elegir alguna. Quizá sea la que estábamos esperando, o quizá sólo sea la menos mala. Por eso prefiero la idea de salir a ofrecerse, decidir primero qué es lo que vamos a hacer, para luego salir al mercado a buscar a un futuro empleador que se interese en nosotros. Finalmente, elegimos esta carrera porque nos gusta, porque esto es lo que queremos hacer de nuestras vidas.

Para un profesional de sistemas existen variedad de caminos a elegir: se puede seguir el camino del desarrollo, del análisis funcional, QA, procesos, etc. Yo quiero dedicar esta nota solamente a aquellos que eligen el desarrollo.

El desarrollador tiene una pasión especial por su tarea; es muy común encontrarse con alguno que se haya pasado el fin de semana probando un nuevo framework o instalando, compilando y configurando una nueva distribución de Linux. También es cierto que en esta categoría entran los que por un motivo u otro no tuvieron la oportunidad de contar con una formación universitaria.

Bien ¿Por dónde empezar? ¿Qué hacer? Es muy probable que a esta altura ya se haya tenido una experiencia de desarrollo informal. Siempre hay disponible un tío, amigo o pariente lejano que necesita un "sistemita" al que acometemos con todo el entusiasmo y con resultados bastante diversos.

También es probable que contemos con un bagaje de "programitas" hechos para nuestro uso personal, y también lo es que contemos con experiencias de trabajos prácticos de la facultad. De cualquier manera es difícil no verse envuelto en una ola de ansiedad a la hora de pensar en una salida laboral. Las empresas tienen otro ritmo, hay responsabilidades que cumplir y, sobre todo, hay que integrarse a una estructura que ya está funcionando y tiene sus propios códigos.

En primer lugar habrá que decidir entre dos grandes grupos. Uno es el mercado empresarial y otro el particular, aunque también podríamos incluir un tercero que es el académico.

Empresas

Comienzo con el sector empresarial. Habrá que salir a llevar Cvs a consultoras impuestas en el mercado de las grandes empresas. En este caso habrá que decidirse (generalmente) entre dos opciones: Java o .Net. Dejo afuera opciones de productos como puede ser SAP ya que generalmente terminan siendo dedicaciones más orientadas a configuración y definiciones funcionales.

Cualquiera de las dos opciones es igualmente buena y nos terminará ubicando en clientes de características similares. Con Java es muy probable que consigamos algún puesto mejor remunerado, aunque hay que tener presente que el desarrollo va a ser más complicado y la curva de aprendizaje es más larga y constante. La opción .Net nos abre la puerta a un mercado más amplio, ya que el desarrollo en esta plataforma no está acotado al mercado enterprise como en el caso de Java. También es una buena opción para aquellos que no se sienten cómodos decidiendo entre una variedad importante de herramientas para cada una de las tareas que deseen realizar. Los que opten por Java deberán elegir qué IDE usar, configurarle los plugins necesarios, amoldarse al framework elegido para visualización, loggin, persistencia, servicios, en fin, todo. Y en cada proyecto irán variando. Con .Net la cosa va más sencilla: Microsoft se ocupa de ofrecer soluciones a todo (y en general, buenas).

Otro punto a tener en cuenta es el soporte para las aplicaciones. En Java hay un esfuerzo importante para lograr la instalación y el correcto funcionamiento de las mismas y mayormente mucho xml. Mucho. Xml para el tomcat, para struts, para hibernate, para spring, para log4j, para ant y, como si esto fuera poco, algunos xml ad-hoc que se crean en el proyecto para configurar vaya uno a saber qué, pero nunca faltan.

Por el lado de .Net, si bien hay xmls, las herramientas se ocupan de administrarlos con un buen nivel de éxito. Una de las cosas que más me molestaron en mi experiencia con la plataforma de MS fue la falta de fuentes. Cuando las aplicaciones tienen una arquitectura compleja con dependencias en varias aplicaciones y librerías, suele pasar que surgen errores difíciles de identificar. Y el hecho de no tener el código fuente hace que uno se quede sin saber bien qué está pasando. En

Juan José Evangelista

Licenciado en Sistemas de la Universidad CAECE y se desempeña actualmente como Coordinador de la oficina de Desarrollo Profesional de Snoop Consulting. También forma parte de la Oficina de Arquitectura de Software de la misma empresa en el rol de Consultor de arquitectura J2EE.

Java es muy raro trabajar con una librería sin contar con sus respectivos fuentes. El caso más claro es el del application server. Uno puede contar con los fuentes de un Jboss, IAS o Websphere, pero nunca se sabe bien qué pasa dentro de la cabeza de un IIS.
En definitiva, y más allá de las diferencias, en cualquiera de las dos tecnologías hay que estar dispuesto a trabajar en grupo, a amoldarse a las distintas políticas de los clientes, a respetar horarios fijos y a seguir procesos preestablecidos, tanto formales como informales.

Particulares

Por otro lado tenemos el mercado particular. Con particular me refiero a una modalidad de trabajo más independiente y con clientes menos definidos. Aquí entra el diseño de sitios web, de pequeños desarrollos para pymes, programas de distribución masiva, juegos, aplicaciones para móviles, etc.
Este es un mercado muy amplio, no es menos remunerativo, pero sí mucho más riesgoso.

También es el más apto para aquellos que no tienen un título universitario.
El protagonista de este momento es PHP. Es un lenguaje muy potente y fue el heredero natural de ASP. Hoy es el punto de partida para el que quiera iniciarse en el desarrollo de aplicaciones web. Es un lenguaje muy flexible que en su versión 5 puede adaptarse a cualquier estilo, pudiendo trabajar de una forma estructurada y clásica o hacerlo orientado a objetos al estilo Java. Cuenta con el apoyo de una comunidad enorme y muy activa y además hay cantidad de proyectos con utilidades de todo tipo.
Por otro lado hay una cantidad importante de aplicaciones completas desarrolladas por la comunidad y distribuidas bajo licencias libres que permiten la comercialización de las mismas sin la necesidad de realizar ningún desarrollo específico.
Los requerimientos de instalación son mínimos y el costo de hosting es muy bajo, transformándolo en la opción ideal para ofrecer un producto para pymes o emprendimientos privados con proyección comercial.
Hace poco tiempo surgió una nueva opción de la mano de Ruby que es "Ruby on Rails". Este es un framework de creación de sitios web muy potente que permite el desarrollo de un proyecto completo en tiempos impensados en otras plataformas. La gran contra de este framework es que toda su potencia se hace evidente cuando tenemos todas las decisiones de nuestro lado, sobre todo en lo que respecta al modelo de datos.
Otra opción importante es Python que es un lenguaje orientado a objetos y que está ganando terreno en el desarrollo de aplicaciones desktop gracias al soporte de librerías de componentes como wxPython o qt.

Lectura Adicional

http://java.sun.com/
www.microsoft.com/net
www.php.net
www.python.org
www.wxpython.org
www.riverbankcomputing.co.uk/pyqt/
www.ruby-lang.org
www.rubyonrails.org

Conclusión

Si bien no intento hacer un análisis exhaustivo, ni tampoco terminante, mi idea es ubicar las diferentes opciones laborales y las tecnologías que las acompañan. En este último punto creo importante aclarar que casi cualquier plataforma es válida para cualquier contexto, pero el papel que cumplen las comunidades y las tendencias de mercado hacen que la balanza se vuelque hacia un lado u otro dependiendo del tipo de problemática.
En nuestra área estamos viviendo un momento muy especial ya que la demanda parece no terminar nunca y las oportunidades y ofertas laborales nos ponen en una posición privilegiada frente a otras profesiones. Nosotros podemos elegir. Hagámoslo con pasión. ●





Nuevos lanzamientos

Kaspersky Lab, líder en desarrollo de sistemas de seguridad informática, anunció el lanzamiento de la nueva versión 6.0 en español de su Antivirus e Internet Security.
Estos dos softwares combinan la detección reactiva con la más reciente tecnología proactiva para proveer una protección sólida contra los programas maliciosos.
Algunas de las nuevas capacidades que ofrecen estos nuevos softwares son el análisis al vuelo de emails y tráfico de Internet, el análisis mediante un antivirus optimizado, la protección contra rootkits, actualizaciones instantáneas de la base de datos y un uso simple y fácil.
Kaspersky Antivirus 6.0 recibió el premio Virus Bulletin 100 por ciento al lograr una detección del 100 por ciento de los virus mientras que al mismo tiempo no generó ningún falso positivo. Asimismo Kaspersky Internet Security 6.0 gano el prestigioso premio "Editor's choice" de la revista alemana Magazine com! ●

Para mayor información sobre los productos dirigirse a: http://www.kaspersky.net.ar/

NOVEDADES DE PRODUCTOS Y EMPRESAS

AVAYA

Nueva solución de comunicaciones IP

Avaya presentó sus nuevos teléfonos IP VPNremote para teletrabajadores, los cuales poseen una mayor interoperabilidad de redes y permite que un mayor número de empresas puedan desplegar comunicaciones con calidad de oficina de manera más rápida y barata.
Los empleados que estén trabajando en localizaciones remotas pueden instalar un teléfono VPNremote IP en su casa con un equipamiento mínimo y con un mínimo de asistencia IT: requiere solamente de una fuente de alimentación eléctrica, un enrutador de banda ancha y un password.
Además de una mayor interoperabilidad, la nueva versión del teléfono Avaya agrega capacidades avanzadas de monitoreo de Calidad de Servicio (QoS) que permiten que un administrador IT externo monitoree a larga distancia la conectividad y calidad de transmisión de voz de las comunicaciones de un usuario remoto. ●

Para más información:
http://www.avaya.com.ar/

Microsoft

ISA SERVER 2006

A pesar de que el nuevo Internet Security and Acceleration (ISA) Server 2006 de Microsoft parece similar a su versión anterior de 2004, el nuevo software trae varias mejoras para sus clientes. Una de las principales es la posibilidad de preservar el ancho de banda mediante la compresión del contenido HTTP, lo que resulta beneficioso si se tiene una lenta conexión WAN entre los clientes y el ISA Server. Además, controla el ancho de banda de todo el tráfico HTTP y HTTPS.
Otro de los beneficios es que el ISA 2006 vigila el flujo de lo recibido evitando los virus y los ataques maliciosos mediante la configuración de un número máximo de pedidos por minuto a la dirección de IP. A casi un mes de la aparición del ISA 2006 aún no se han detectado problemas, lo que la convierte en la mejor versión que salió al mercado de la mano de Microsoft. ●

Para más información:
http://www.microsoft.com/isaserver/prodinfo/default.mspx



Microsoft

Cisco y Microsoft
en pos de la interoperabilidad

Dos de las empresas más importantes a nivel mundial anunciaron el lanzamiento de una arquitectura de seguridad que permite a los clientes y socios instalar las infraestructuras Network Admission Control - NAC (Control de Adminisión a la Red) de Cisco y Network Access Protection – NAP (Protección de Acceso a la red) de Microsoft.

Un documento técnico distribuido por las dos compañías describe cómo estos dos protocolos interoperan para reforzar las políticas de seguridad y asesoría en cuanto a un plan de cuidados de la salud de la red. El documento describe la arquitectura y provee detalles referidos a cómo integrar las capacidades de seguridad embebida de la infraestructura de la red Cisco con la cual trabajan tanto Windows Vista como la futura versión de Windows Server (cuyo nombre en código es “Longhorn”).

Aun cuando ambas empresas van a continuar distribuyendo sus soluciones individuales, los clientes tendrán ahora una opción con el Cisco NAC, el Microsoft NAP, o la solución interoperable de ambas firmas. ●

Para más información:
http://www.cisco.com/go/nac
http://www.microsoft.com/nap

BREVES

A 50 años del primer "disco duro"

IBM conmemora 50 años del lanzamiento del primer disco de almacenamiento de la empresa y de la historia de la tecnología. La presentación del RAMAC Computer (Random Access Memory Accounting Machine), el día 13 de septiembre de 1956, inició una revolución tecnológica. Cada equipo tenía 5 Megabytes de capacidad de almacenamiento, mientras que en la actualidad un disco puede superar los 200 Terabytes (200 millones de megabytes). El RAMAC era tan grande que para trasladarlo hacía falta un camión. Sin embargo, su capacidad de almacenamiento era infinitamente menor a la que posee hoy un CD que almacena más de 700 MB y cuesta $2, en tanto que este primer dispositivo costaba U$50.000.





Para más información visite *www.ibm.com/ar*

Google se agranda

El conocido buscador anunció el lanzamiento de un paquete de programas diseñados especialmente para empresas llamado Google Apps for Your Domain. Este software, que ya está disponible para ser bajado en forma gratuita es un set de soluciones que incluye la administración de GMail, Google Talk (mensajería instantánea), Google Calendar (agenda compartida) y Google Page Creator (páginas web).

Si a esto le sumamos el hecho de que desde hace un tiempo está disponible el Google Writely como un procesador de textos online podemos estar frente a un nuevo competido del Office de Microsoft. Si quiere más información acerca de los dos proyectos de Google, ingrese a:
http://www.google.com/a ó *http://www.writely.com/*

Incident Response & Computer Forensics

Escrito por personal del FBI, Kevin Mandia, Chris Prosise y Matt Pepe, la segunda edición del best seller *Incident Responce & Computer Forensics* ofrece una mirada sobre los aspectos legales, los procedimientos y los pasos técnicos de la computación forense. Además incluye nuevos capítulos sobre el análisis forense y sus soluciones y un detallado estudio de un caso real. Este libro revelador muestra cómo poder contrarrestar y vencer los ataques de los hackers y las actividades maliciosas.

Con la cobertura de los sistemas operativos Windows y Unix y de recursos de plataformas no específicas como los servicios de la Web y routers, este libro de 544 páginas devela las herramientas necesarias para facilitar la investigación en la computación forense.

Para más información:
www.amazon.com

Microsoft y Citrix extienden su relación comercial

Microsoft y Citrix anunciaron que planean extender su relación comercial existente mediante la distribución integral de aplicaciones para facilitar, hacer más seguro y más económico el uso de computadoras para los empleados de las sucursales. Las compañías colaborarán en el desarrollo y la comercialización de un nuevo dispositivo multifuncional de Citrix basado en el sistema operativo Windows Server, en el ISA Server y en el Citrix WANScaler.

La solución conjunta planeada representa un nuevo tipo de dispositivo para sucursales que combina tecnologías de optimización avanzadas para redes de área extensa (WAN) con servicios consolidados para sucursales en un dispositivo multifuncional. Esta combinación ayuda a resolver un problema mayor para los clientes quienes antiguamente mediante las herramientas tradicionales para redes o productos puntuales individuales no podían.

Para más información:
http://www.citrix.com
http://www.microsoft.com

Humor - *Por Severi*

